École des Hautes Études en Sciences Sociales

Ecole doctorale 465 – Economie Panthéon Sorbonne

Paris-Jourdan Sciences Economiques

PhD Thesis

Field: Economics - Speciality: Analysis and Policy in Economics

Prepared and defended at the Paris School of Economics on **April 22nd, 2024***, by*

Heim Arthur

Social investment & the changing face of poverty

Essays on the design and evaluation of family and social policies in France

PhD advisor: Marc Gurgand

Reviewers :			
	1	Anne Boring	Erasmus School of Economics, Rotterdam
	2	Camille Terrier	Queen Mary University London

Jury:			
	1	Anne Boring	Erasmus School of Economics, Rotterdam
	2	Marc Gurgand	CNRS & Paris School of Economics
	3	Karen Macours	INRA & Paris School of Economics
	4	Rafael Lalive	University of Lausanne
	5	Camille Terrier	Queen Mary University London

École des Hautes Études en Sciences Sociales

Ecole doctorale 465 – Economie Panthéon Sorbonne

Paris-Jourdan Sciences Economiques

Doctorat

Discipline: Sciences Économiques - Spécialité: Analyse et Politique Économiques

Heim Arthur

Investissement social & nouvelles formes de pauvreté

Essais sur la mise en oeuvre et l'évaluation de politiques sociales et familiales en France

Thèse dirigée par : Marc Gurgand

Date de soutenance : le 22 Avril 2024

Rapporteuses :	1	Anne Boring	Erasmus School of Economics, Rotterdam
	2	Camille Terrier	Queen Mary University London

Jury:	1	Anne Boring	Erasmus School of Economics, Rotterdam
	2	Marc Gurgand	CNRS & Paris School of Economics
	3	Karen Macours	INRA & Paris School of Economics
	4	Rafael Lalive	University of Lausanne
	5	Camille Terrier	Queen Mary University London

Contents

Avant-propos/Forword

Résumés

Summary

This thesis explores early childcare and activation policies, fundamental within the social investment paradigm, through two large field experiments in France, supported by the National Family Allowance Fund.

In the first chapter, with Julien Combe, we consider access to daycare as a matching problem. We propose market design models to define assignment mechanisms and analyse the consequences of design choices in a field experiment. The problem is akin to school choice, but specific constraints affect the definition and scope of stable matchings. Our algorithms provide Student Optimal Fair Assignments (SOFA) in different versions of the problem. Our analysis focuses on the Matthew effect, demonstrating how design and policy choices influence it. Our tools promote fairness and transparency in assignment processes.

Chapters 2 and 3 analyse data from an intensive experimental programme aimed at low-income single-parent families in France, implemented from 2018 to 2022. In Chapter 2, I analyse the effects on labour market participation and poverty, and how wrong we would have been not to use a randomised controlled trial. The analyses reveal initially negative effects that diminish over time. Participants have higher employment rates than other comparison groups, but this difference is entirely due to selection bias. This bias is so strong that estimates using the next best identification strategy - modern doubly robust differences-in-differences - fail to include experimental estimates within confidence intervals. Overall, the programme has no average effect on labour market participation and poverty after the end of the training. There are heterogeneous treatment effects by number of children at baseline.

In Chapter 3, with Alexandra Galitzine, we challenge the narrative of "making work pay" for single-parent families in France. The 2019 reform of in-work benefits (*Prime d'activité*) was adopted contemporaneously with this programme. The intervention directly provided individualized and detailed information on the socio-fiscal system in a year-long support programme, likely to have further reduced various barriers to employment. We use this experiment to measure low-income single-parent families' reactions to incentives after the reform. Our primary contribution lies in estimating counterfactual distributions using experimental assignment variations. We find high labour income elasticities for participants, indicating significant disincentives to employment and increased in-work poverty. The programme's effects on family structure vary based on the number of children, highlighting the complex interplay between policy incentives and poverty dynamics. We coined the term "Assistaxation" to describe the phenomenon of heavily taxing the economic, physical, and mental resources of those accessing public assistance, leaving them with little means to escape.

- **Keywords**: inequalities, public policies, poverty

Résumé de la thèse

Cette thèse explore les politiques d'accueil du jeune enfant et et les politiques d'activation, fondamentales dans le paradigme de l'investissement social, à travers deux grandes expérimentations de terrain en France, soutenues par la Caisse nationale des allocations familiales.

Dans le premier chapitre, avec Julien Combe, nous abordons l'accès en crèche comme un problème d'appariement. Nous proposons des modèles de *market design* pour définir les mécanismes d'attribution et analyser les conséquences des choix de conception dans une expérimentation de terrain. Le problème est similaire au choix d'école, mais des contraintes spécifiques affectent la définition et la gamme des appariements stables. Nos algorithmes fournissent des attributions *équitables* optimales pour les familles (SOFA) dans les différentes versions du problème. Notre analyse porte sur l'effet Matthieu, montrant comment les choix de conception et de politique l'influencent. Nos outils favorisent l'équité et la transparence dans les processus d'attribution.

Les chapitres 2 et 3 analysent les données d'un programme expérimental intensif d'accompagnement global visant les familles monoparentales en situation de précarité en France, mis en œuvre de 2018 à 2022. Dans le chapitre 2, j'analyse les effets sur la participation au marché du travail et la pauvreté, et à quel point nous aurions eu tort de ne pas avoir utilisé un essai randomisé contrôlé. Les analysent révèlent des effets initialement négatifs qui s'estompent avec le temps. Les participants ont des taux d'emploi plus élevés que les autres groupes de comparaison, mais cette différence est entièrement due au biais de sélection. Ce dernier est si fort que les estimations utilisant la prochaine meilleure stratégie d'identification - les différences en différences modernes et doublement robustes - ne parviennent pas à inclure les estimations expérimentales dans les intervalles de confiance. Au final, le programme n'a pas d'effet moyen sur la participation au marché du travail et le taux de pauvreté une fois l'accompagnement terminé. Il y a en revanche des effets hétérogènes en fonction du nombre d'enfants au départ.

Dans le chapitre 3, avec Alexandra Galitzine, nous remettons en question le discours de "*rendre le travail payant*" pour les familles monoparentales en France. La réforme de 2019 de la prime d'activité a été adoptée dans même temporalité que ce programme. L'intervention a directement fourni des informations individualisées et détaillées sur le système socio-fiscal dans un accompagnement d'un an, susceptible d'avoir réduit, davantage encore, les divers freins à l'emploi. Nous utilisons cette expérience pour mesurer les réactions des familles monoparentales pauvres aux incitations après la réforme. Notre contribution principale réside dans l'estimation des distributions contre-factuelles en utilisant les variations expérimentales des affectations. Nous trouvons des élasticités de revenu du travail élevées pour les participantes, indiquant des désincitations importantes à l'emploi et une pauvreté laborieuse accrue. Les effets du programme sur la structure familiale varient en fonction du nombre d'enfants, soulignant l'interaction complexe entre les incitations politiques et les dynamiques de pauvreté. Nous avons proposé le terme "*Assistaxation*" pour désigner ce phénomène consistant à taxer massivement les ressources économiques, physiques et mentales des personnes ayant recours à l'aide publique, leur laissant au passage peu de moyen de s'en extraire.

- **Keywords**: inégalités, politiques publiques, pauvreté

Remerciements

En écrivant les derniers mots de cette thèse, je réalise que ce projet a réellement commencé lorsqu'en 2016, j'ai été recruté à France stratégie pour travailler sur l'investissement social. À partir de ce moment, je crois que je n'ai cessé d'être intrigué, contrarié, contredit, excité, fasciné,… par ce sujet qui ne se tarissait pas; par toutes les personnes qu'il m'a permis de rencontrer et auprès de qui je n'ai fait qu'apprendre.

Non. Cette thèse a débuté dès mon premier emploi, après mon master *Public policy and Development* en 2013. Je rejoins alors le Conseil national d'évaluation du système scolaire (Cnesco) nouvellement installé, et dans lequel Marc Gurgand siège au conseil scientifique. C'est à travers cette expérience qu'a débuté cette collaboration. Je parlais alors déjà de retourner en thèse après quelques années. Je souhaite donc exprimer ma plus profonde gratitude à ce chercheur qui n'a cessé de m'inspirer; ce mentor qui n'a fait que me soutenir, et ce directeur de thèse, que je recommanderais à tout étudiant ou étudiante. Les interactions avec lui tout au long de ce doctorat m'ont incité à la plus haute rigueur académique et m'ont appris à pousser mon niveau d'exigence dans l'analyse économique toujours plus loin. Je lui suis également reconnaissant pour avoir su me soutenir dans les moments de creux, avec une bienveillance que je souhaite souligner. Avoir été son doctorant est un honneur, et j'espère être et continuer à être à la hauteur de sa supervision.

Je tiens ensuite à remercier Karen Macours pour ses excellents conseils, son soutien à travers le comité de thèse. J'ai trouvé chacune de nos discussions intellectuellement stimulantes ainsi qu'inspiratrices de nouvelles idées. Je tiens également à exprimer ma profonde gratitude pour les discussions que nous avons eu en dehors du comité et qui m'ont souvent permi d'élargir mes perspectives.

Je suis également très reconnaissant et honoré d'avoir Anne Boring, Raphael Lalive et Camille Terrier pour jury, et tiens à les remercier d'avoir accepté de lire ma thèse. Les commentaires lors de la pré-soutenance m'ont permis non pas d'améliorer mais de transformer mes travaux.

Nombreux sont les chercheurs qui ont contribué à me faire progresser en tant qu'économiste, mais peu m'ont autant appris que Julien Combe. Je crois qu'il est impossible d'estimer le rendement de ce premier verre de vin partagé lors d'une soirée "Wine & Cheese" de Synapse[1]. ISAJE n'aurait jamais été ce projet sans son implication sans faille, et ce que je peux apprendre à ses côtés ne semble borné. Récemment habilité à diriger des recherches, je souhaite à n'importe quel futur étudiant ou étudiante d'avoir le privilège d'être son doctorant.

Je tiens également à remercier Denis Fougère auprès de qui j'ai énormément appris lors de nos travaux sur le rendement de l'investissement social. Avoir l'honneur d'enseigner l'évaluation des politiques publiques à Sciences Po à ses côtés a été pour moi une expérience extrêmement enrichissante et valorisante. J'éprouve une profonde gratitude pour l'amitié et la confiance manifestées à travers toutes ces années.

Cette thèse n'aurait pas cette contenance sans les dizaines d'heures de discussions, relectures, sessions de brainstorming, flipcharts, etc., avec Alexandra Galitzine, également co-autrice du chapitre III. Merci à elle pour sa rigueur, son ouverture, son engagement, ses apports politistes et féministes qui ont beaucoup nourri cette thèse. Je crois que je n'aurais pu espérer une collaboration plus enrichissante et il me tarde de lancer ensemble d'autres projets.

Je remercie également du fond du cœur Laudine Carbuccia pour son amitié, sa passion et son soutien sans faille, ainsi que pour les innombrables discussions qui nous amènent aujourd'hui vers de nouvelles collaborations. Je remercie également Nisrynne pour ses relectures, discussions et son aide dans la préparation de la soutenance.

[1] l'association étudiante de Paris School of Economics.

Ma carrière est marquée par des personnes qui ont dû prendre la décision de me faire confiance, et sans qui je ne serais pas ici aujourd'hui. J'adresse mes remerciements et pensées sincères à Eric Maurin, qui a accepté de superviser mes travaux de Master puis m'a aidé dans ma recherche d'emploi. Je remercie également Nathalie Mons et Jean-François Chesné pour leur confiance au Cnesco. Ces remerciement me font également avoir une pensée émue pour François Dumas, son ancien Secrétaire général. François fait pourtant partie de ces personnes qui m'ont accordé une confiance – probablement déraisonnable – mais qui m'ont permis de nourrir cette curiosité insatiable et de devenir le chercheur que je suis. François fait malheureusement partie de ces regards qui manqueront lourdement lors de cette soutenance, ayant croisé la faucheuse avant l'arrière-saison.

Je remercie également Sandrine Dauphin et Marine Boisson-Cohen pour m'avoir lancé sur l'investissement social et laissé libre d'explorer la question sans limite ; Jean Pisani-Ferry et Selma Mahfouz de m'avoir recruté et Fabrice Langlart pour ses relectures et discussions techniques et stimulantes. L'expérience à France stratégie a été pour moi un moment d'apprentissage et de remises en question importantes, notamment grâce aux discussions riches et animées avec Bruno Palier, Julien Damon, Cyprien Avenel, Christophe Fourel, Nicolas Duvoux. Elle m'a aussi permi de travailler aux côtés de collègues aussi brillants que chaleureux. Je remercie tout particulièrement Johanna Barasz, Marine De Montaignac, Mathilde Viennot, Clément Dherbécourt, Daniel Agacinski, Pierre-Yves Cusset, et tous les autres pour ces années riches de souvenirs. Merci enfin à Gautier Maigne de m'avoir donné les moyens de faire des travaux aussi ambitieux et différents, de m'avoir accordé sa confiance pour être conseiller scientifique et impliqué dans l'évaluation de la stratégie pauvreté, de m'avoir soutenu tout au long de ma carrière. Je remercie également Louis Schweitzer de m'avoir donné autant la possibilité de nourrir les travaux du comité d'évaluation, notamment avec Élise Huilery et Sébastien Grobon. Je suis honnôré et fier d'avoir pu travailler et apprendre avec des chercheurs et chercheuses aussi inspirantes.

Les travaux de cette thèse n'auraient pas été possibles sans les soutiens institutionnels et financiers de la Caisse nationale des allocations familiales, mais surtout la confiance accordée par mes responsables. Je tiens à exprimer toute ma reconnaissance à Bernard Tapie, Florence Thibault, Jeanne Moeneclaey, Virginie Gimbert, Lucie Gonzales, Daniel Lenoir, Vincent Mazauric et Nicolas Grivel pour avoir cru en moi, malgré les difficultés et l'ampleur des chantiers. Je remercie toute l'équipe de la direction des statistiques des études et de la recherche de la Cnaf auprès de qui j'ai passé une bonne partie de ces 6 années. J'adresse une attention toute particulière à Saad Loutfi dont l'intelligence, le professionalisme et la bonne humeur n'ont fait qu'enrichir et favoriser le succès d'ISAJE. Merci également à Cécile Ensellem, Anne Unterreiner, Benoit Céroux, Clémence Helfter, Sephora Besançon et Chloé Chasagnac, pour avoir notamment - et notoirement - mis de la bonne humeur dans mes longues journées de travail.

Je souhaite également remercier Pierre Boyer, Antoine Bozio, Julien Grenet, Delphine Roy et l'Institut des Politiques Publiques (IPP) - Paris School of Economics pour nous avoir soutenu et fourni des assitants de recherches pour la recherche ISAJE. Je remercie tout particulièrement Paul-Emmanuel Chouc et Vincent Verger (IPP), pour leur travail rigoureux et leur investissement sur diverses vagues d'affectation, Agathe Eupherte, Raja Ahmed Taleb et Marion Goglio pour leur travail remarquable en tant qu'assistants de recherche. I also thank the members of the ISAJE scientific committee for helpful guidance through this work and especially Orla Doyle and Sylvana Cote.

Je remercie également le Conseil départemental de Meurthe-Et-Moselle, l'équipe Reliance et toutes les personnes impliquées dans ce projet. Je tiens en particulier à remercier Gabriel André pour son audace et son engagement à changer les conditions de vies des moins fortunés.

Je remercie également les directions petite-enfance et élus des territoires impliqués dans ISAJE sans qui nous n'aurions pu apprendre tant.

Je tiens également à remercier toutes celles et ceux qui m'ont fait bénéfier de leurs remarques et commentaires lors de présentations, séminaires, relectures ou simples discussions. Beaucoup sont nommés en début de chapitres mais la liste et loin d'être exhaustive.

Enfin, un grand merci à mes amis de longue date pour leur soutien indéfectible tout au long de ce parcours. Leur présence et leur soutien ont été une source de réconfort et de motivation inestimable. Je suis également reconnaissant envers ma famille pour son amour et son soutien à travers les années. Un grand merci également à mes colocataires, ainsi qu'à mes amants et amantes, passées et présentes, pour leur soutien et leur présence à mes côtés tout au long de ce voyage.

"*Many of us chose economics because, ultimately, we thought science could be leveraged to make a positive change in the world. There are many different paths to get there. Scientists design general frames, engineers turn them into relevant machinery, and plumbers finally make them work in a complicated, messy policy environment. [...] A feature unique to economics is that scientists, engineers, and plumbers all talk to each other (and in fact are often talking to themselves — the same economist wearing different hats)*" — Esther Duflo (2017), *The Economist as Plumber*.

Investissement social et nouvelles formes de pauvreté: Introduction générale

Cette thèse s'intitule *Investissement social et nouvelles formes de pauvreté* et s'intéresse en particulier au problème des inégalités d'accès aux modes d'accueils formels et aux politiques de lutte contre la pauvreté à destination des familles monoparentales. Ces travaux ont été menés à la Caisse nationale des allocations familiales dans le cadre d'un contrat CIFRE entre 2019 et 2022, puis d'un CDI jusqu'à la fin de cette thèse. Les trois années qui ont précédé ce projet ont été consacré à l'investissement social dans une mission jointe entre France stratégie et la Cnaf, et c'est au cours de cette première vie professionnelle que ce projet s'est consolidé.

L'investissement social est une approche visant à reconfigurer l'intervention sociale et la protection sociale pour mieux accompagner les individus dans leurs trajectoires de vie[2] en réponse à l'évolution des besoins sociaux. Cette approche met l'accent sur la nécessité d'intervenir en amont des situations, notamment dès la petite enfance, afin d'anticiper et de prévenir les risques. Bruno Palier (2014) propose ainsi cette définition simple: "*l'investissement social invite à préparer l'avenir pour avoir moins à réparer*".

Il apparait dans une période où les modèles d'États providence sont bousculés par des mutations profondes des sociétés générant de nouveaux défis : transformations de la famille avec l'accroissement des divorces et la diversification des formes de ménages ; transformations des relations sociales avec la baisse des grands collectifs d'actions, syndicats et clergés, les fortes aspirations à l'émancipation et la reconnaissance pour l'égalité des femmes, des personnes racisées, des personnes LGBTQIA+, etc; transformations de l'économie et des institutions avec la mondialisation, la construction européen, les progrès technologiques,…

Les mutations sociétales ont profondément perturbé les mécanismes traditionnels de l'État social et mis en lumière ses limites, engendrant une crise tripartite : financière, d'efficacité et de légitimité. Dans un contexte de contraintes budgétaires, la prise en charge des risques sociaux traditionnels tels que le chômage, la santé et les retraites devient de plus en plus problématique, tandis que les nouveaux risques ne sont que partiellement couverts : la pauvreté infantile, les difficultés d'insertion professionnelle des jeunes et des moins qualifiés, les défis liés à l'articulation entre vie familiale et vie professionnelle, ainsi que la perte d'autonomie. Les enjeux de soutenabilité ne se posent plus qu'en terme de dette publique et la gestion des conséquences de l'anthropocène mettent, davantage encore, les sociétés sous tension[3].

L'investissement social a intégré l'agenda européen en 2013 avec la présentation du « Paquet Investissement Social » (PIS) qui forme un cadre pour la réorientation des politiques sociales des États membres autour de la notion d'investissement social tout au long de la vie. Une définition a alors été proposée pour les investissements sociaux qui « consistent à investir dans les personnes en adoptant des mesures pour renforcer leurs compétences et leurs capacités et leur permettre de participer pleinement au monde du travail et à la société. Les domaines prioritaires sont l'éducation, les services de garde d'enfants de qualité, les soins de santé, la formation, l'aide à la recherche d'emploi et la réinsertion » (Européenne 2015). Cette conception renouvelée de la protection sociale — *lato sensu* incluant politiques de l'emploi, politiques éducatives — repose sur l'idée que l'argent dépensé aujourd'hui pourrait rapporter à un moment donné dans le futur (caractéristiques de l'investissement) ; le rendement attendu incluant en particulier des dépenses évitées de protection sociale (Palier 2014).

[2] Voir notamment le livre Avenel et al. (2017) reprenant les ensignements d'un cycle de séminaire sur la définition d'une stratégie d'investissement sociale pour la France.

[3] Cette thèse n'adresse néanmoins pas du tout les enjeux de soutenabilité écologique.

Cette définition proposée par la Commission Européenne ne fait pas nécessairement l'unanimité car l'investissement social a trait également aux théories de la justice sociale et se voit parfois attribuer d'autres finalités, notamment celles de concilier des objectifs économiques avec un objectif social (maintien de la cohésion sociale par la lutte contre les inégalités) et démocratique (permettre l'exercice d'une citoyenneté active). Elle met au moins en avant les deux principes qui constituent le socle consensuel de l'investissement social : l'investissement dans le capital humain des individus et l'objectif de participation au marché du travail (Deeming and Smyth 2015; Garritzmann, Häusermann, and Palier 2022).

Dans cette introduction, je propose d'abord de revenir sur les transformations de la société brièevement évoqués pour clarifier ce que je nomme ici les *nouvelles formes de pauvreté*. Dans la section II, je discute rapidement de quelques formes d'institutionnalisation de l'approche investissement social en France, motivant le choix de ce cadre analytique pour l'annonce de la problématique dans la Section III. La section IV présente le cadre d'analyse et les deux grands projets menées pour répondre à la problématique. Enfin, je discute les contributions de chaque chapitre dans les trois dernières section: La section V synthétise les travaux sur les mécanismes d'affectations des places en crèche et leurs effets sur les inégalités ; la Section VI résume les effets du programme Reliance, un programme d'accompagnement intensif à destination des cheffes de famille monoparentales au RSA longue durée et insiste sur l'importance des méthodes d'évaluation. Enfin, la Section VII récapitule l'analyse des incitations monétaires au travail et leurs effets sur les familles accompagnées dans le programme Reliance.

I Les transformations de la société et l'émergence de nouveaux risques

Cette année, le prix de la Banque de Suède en sciences économiques en mémoire d'Alfred Nobel a été décerné à Claudia Goldin pour avoir avancé la compréhension de la participation des femmes au marché du travail. Elle est ainsi la troisième femme économiste à recevoir ce prix, 2 ans après Esther Duflo[4] pour ses travaux sur l'approche expérimentale pour réduire la pauvreté. Dans une célèbre contribution Claudia Goldin (2006-May) analyse les facteurs menant à ce qu'elle appelle la "révolution discrète[5]" qui a transformé l'emploi, l'éducation et la famille des femmes.

I.1 La révolution du genre en économie

Les travaux de Goldin fournissent des éléments essentiels pour comprendre les phénomènes qui nous préoccupent, notamment les divergences significatives dans les parcours éducatifs des femmes et des hommes au cours des cinquante dernières années. Ces travaux mettent également en lumière l'émergence de groupes de population qui se retrouvent plus fréquemment en situation de précarité, parmi lesquels les familles monoparentales occupent une place importante.

Les prémisses de la révolution Pour mener à cette révolution, Claudia Goldin (2006-May) identifie trois changements touchant les femmes:

- 1er changement: **L'horizon temporel**. Si la participation au marché du travail a vocation a être intermittents ou plus permanents, les choix d'éducation ne sont pas les mêmes.
- 2ème changement: **L'identité**. À savoir d'un côté, si une femme considère que son identité est liée à son rôle d'épouse, de mère, ou si elle s'identifie davantage à son occupation, son niveau de qualification,… et de l'autre[6], l'acceptation, la valorisation ou dévalorisation de l'expression de ces identités par sa ou son partenaire, ses proches, les entreprises, la société, à travers les normes sociales, les politiques publiques et les institutions des Etats providence.
- 3ème changement: **Le processus de décision**, qui peut être commun aux membres du foyer, coopératif ou stratégique, et de plus en plus dépendre d'options extérieures ; que ce soit sur le marché du travail ou sur les marchés plus informels des rencontres et relations sociales.

[4] Prix partagé avec Abhijit Banerjee et Michael Kremer.
[5] Le titre de l'article est *The Quiet Revolution That Transformed Women's Employment, Education, and Family*
[6] Claudia Goldin n'évoque qu'en filigrane cette seconde partie.

La révolution discrète trouve ses prémisses dans la génération du baby boom. Les mères des enfants de cette génération ont pour beaucoup travaillé pendant la guerre mais se sont re-spécialisés dans la production domestique conformément aux normes patriarcales de l'époque. Cependant, l'accès à des technologies et les changements des modes de consommation ont libéré du temps (aux femmes et aux hommes, voir infra) à la fois pour travailler et/ou s'occuper de l'éducation des enfants (Greenwood, Seshadri, and Yorukoglu 2005; Bose, Jain, and Walker 2022) et des changements législatifs facilitant les temps partiels et horaires atypiques ont permis aux femmes de travailler davantage.

Dans les années 1950, la demande de main-d'œuvre a fortement augmenté et, pour les femmes, elle a traversé une fonction d'offre de main-d'œuvre relativement stable et plutôt élastique *i.e.* à mesure que la demande s'intensifiait, davantage de femmes sont entrées sur le marché du travail. Mais les emplois disponibles restaient limités, non seulement du fait de leurs compétences et niveaux de diplômes mais aussi en raison de normes de genre au sein des entreprises, limitant les perspectives de progression (Toossi 2002). Les générations suivantes ont extrapolé sur cette base pour former des attentes plus précises quant à leur avenir. C'est l'évolution de l'horizon. Ce faisant, une part de plus en plus grande de femmes[7] s'est mise à investir dans son éducation. Elles ont progressivement pu accéder à des emplois mieux rémunérés, plus stables, réduisant notamment l'écart de rémunération entre les femmes et les hommes (voir notamment Figure 1 et 3 ci-après).

Pour la sociologue McLanahan (2004a), le féminisme de la 2ème vague (des années 1960-1970) a fortement influencé le changement d'identité individuelle et sociale des femmes portant la voix du combat pour l'indépendance et l'égalité entre les genres, notamment dans l'accès à l'éducation et à des emplois de qualité. Surtout, ces mouvements ont fortement critiqué la spécialisation des rôles genrés dans la famille, en produisant de nouvelles normes plus égalitaires pour les mariages et en s'opposant à la stigmatisation des mamans-solos. En fournissant "*une identité autre que celle d' "épouse" et de "mère", [le féminimsme] les a encouragées à s'investir dans leur éducation et leur carrière*" (P. 617).

La dernière évolution portant sur le processus de décision s'analyse en terme de dynamiques de pouvoir qui se manifeste au sein du foyer et dans l'évolution des normes et lois. Cette évolution n'est pas qu'une conception abstraite et théorique du pouvoir, ce sont des changements très concrets dans le droit de prendre des décisions en autonomie. Rappelons qu'en France, les femmes ne pouvaient s'inscrire à l'université sans l'accord de leur père ou mari avant 1938, élire leurs représentant·es avant 1944, gérer leurs propres biens et travailler sans l'accord de leur mari avant 1965, user de contraception avant 1967, avorter avant 1975[8],… D'autres groupes aux droits civiques réduits ont vu leurs conditions d'existence s'améliorer et leurs droits progresser, généralement dans des affrontements violents et une résistance forte d'une partie de l'opinion et des institutions. Le mouvement pour les droits civiques aux États-Unis est entaché d'événements sanglants et de victoires en faveur d'une plus grande égalité. Des révoltes de Stonewalls en 1969 naissent les mouvements pour les droits des LGBTQIA et une lente et précaire convergence vers l'égalité de droit[9]. L'évolution des droits s'accompagnent également de chocs technologiques et notamment des méthodes contraceptives et abortives.

Le rôle crucial de la contraception et de l'avortement L'accès à la contraception mais surtout la légalisation de l'avortement a permis aux femmes de reporter ou abandonner l'idée d'avoir des enfants pour investir dans leur éducation et leur carrière, de retarder l'âge du mariage, d'être plus sélectives dans le choix de leurs partenaires, favorisant des unions avec des personnes plus à même de fournir un soutien émotionnel et de participer à l'éducation des enfants (McLanahan 2004a). Les analyses économétriques des effets à long terme de l'accès aux technologies permettant un meilleur contrôle des décisions de fertilité (contraception, tests de grossesse, avortement mais aussi procréation médicalement assistée) soutiennent sans ambiguïté cette thèse. Les recherches montrent que l'accès a l'avortement a fortement contribué au report des décisions maritales et de fertilité et baissé la natalité pour certains groupes, notamment celles qui ont opté pour une éducation plus longue (Claudia Goldin and Katz 2002; Myers 2017). L'effet de la pilule s'analyse en deux temps: à court terme, une partie des cohortes affectées utilise la pilule pour poursuivre des études supérieure et retarder leur décisions de fertilité quand, en même temps, la part d'enfants nés

[7] Nous revenons par la suite sur le rôle important de qui participe à la révolution silencieuse.

[8] L'avortement a même été puni de peine de mort pendant le régime de Vichy et des exécutions ont eu lieu.

[9] Aux États-Unis, les discriminations à l'emploi en raison du genre et de l'orientation sexuelle ne sont illégales au niveau fédéral que depuis un Arrêt de 2020 de la Cour suprême dans l'affaire *Bostock v. Clayton County*, dans lequel la Cour a estimé que la discrimination à l'encontre des personnes LGBTQ est une forme de discrimination sexuelle interdite en vertu du titre VII de la loi de 1964 sur les droits civils.

hors mariage et dans des familles monoparentales a augmenté et le taux de mariage précipités (*shotgun marriage*) s'est fortement réduit (Stevenson and Wolfers 2007). À plus long terme, l'accès à la pilule a augmenté la part d'enfants dont la mère est diplômée du supérieur et baissé la part d'enfant de parents divorcés (Elizabeth Oltmans Ananat and Hungerman 2012; Cheng 2022). À long terme, celles affectées par l'introduction de la contraception et la légalisation de l'avortement ont en moyenne des enfants plus tardivement et participent davantage au marché du travail. Elles ont également une probabilité plus élevée de ne jamais s'être mariée et pour celles qui se marient, une probabilité plus faible de divorcer.

L'accès à la contraception et à l'avortement ont également réduit la pauvreté Gruber, Levine, and Staiger (1999) se demandent qui sont les "enfants marginaux" i.e. quels enfants ne sont pas nés du fait de l'avortement (et donc quels auraient été leurs destinées) en comparant les cohortes affectées ou non par la légalisation de l'avortement suite à la décision *Roe v. Wade* de la cour suprême en 1973. Leurs estimations impliquent que ces enfants auraient une probabilité 70 % plus élevée de vivre dans une famille monoparentale, 40 % plus élevée de vivre en pauvreté, 35 % plus élevée de mourir pendant l'enfance et 50 % plus élevée de recevoir de l'aide sociale. Bitler and Zavodny (2002) montrent que la légalisation de l'avortement a diminué les maltraitances envers les enfants. John J. Donohue and Levitt (2001) ont amorcé un débat académique important autour de l'effet causal de la légalisation de l'avortement sur la baisse de la criminalité dans les années 1990. Dans ce papier, les auteurs comparent les cohortes des États affectés tôt par *Roe vs. Wade* avec ceux qui légalisent plus tard et mesurent des effets forts sur la criminalité. Cependant, cette première étude a été critiquée pour la qualité de ses données (Foote and Goetz 2008 ; John J. Donohue III and Levitt 2008), la puissance statistique et la non prise en compte d'autres facteurs importants comme la déségrégation raciale et sociale (Joyce 2001; John J. Donohue and Levitt 2004 ; Joyce 2009). Malgré les controverses, John J. Donohue and Levitt (2019) confirment et maintiennent leurs résultats 20 ans plus tard (et avec 20 ans de données supplémentaires).

Les chocs technologiques ont donc affecté fortement la capacité de certaines femmes de reporter leur décisions de fertilité, d'augmenter leur qualifications mais tous ces travaux mettent en avant un effet de sélection fort (Elizabeth Oltmans Ananat et al. 2009), pouvant s'expliquer par des profils aux normes et préférences différentes mais aussi des difficultés d'accès à ces technologies (B. P. Brown et al. 2020). Les mouvements anti-avortement ont par ailleurs été très virulents; Medoff (2003) montre que certaines actions contre les centres de planning familial et de cliniques d'avortement ont contribué à réduire l'offre de soin sans affecter la demande, réduisant l'accès à l'avortement aux femmes les plus précaires en premier lieu.

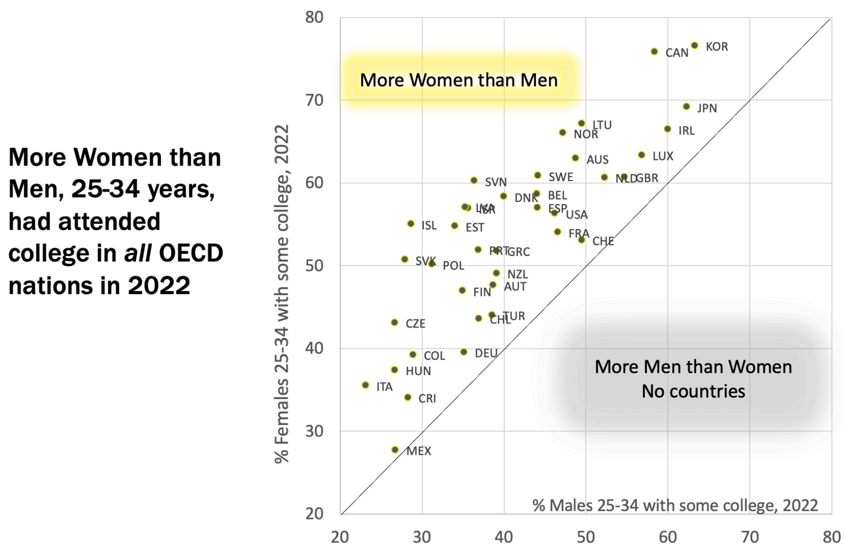
Les inégalités dans l'accès à la contraception et l'avortement affectent fortement le bien-être de la mère. Dans une comparaison de données longitudinales sur 40 pays, Martínez et al. (2022) montrent que l'accès à l'avortement est fortement prédictif d'un moindre syndrôme dépressif post-partum. Pouvoir davantage choisir d'avoir ou non un enfant en ayant recours à un avortement est déterminant pour l'éducation et l'activité des mères mais aussi de l'investissement dans l'éducation des enfants. Les décisions de fertilité et de recours à l'avortement sont également sensible aux résultats d'élection ; Dahl, Lu, and Mullins (2022) montre que celle de Donald Trump a provoqué des changements très importants dans les naissances et avortements, hétérogènes entre localités et alignées avec les préférences politiques. Gonzalez, Guirola, and Zapater (2022) obtient des résultats similaire en Espagne lors des élections de 2004, gagnée par les sociaux-démocrates.

Les femmes sont en moyenne plus diplômées que les hommes La révolution de Claudia Goldin désigne donc la hausse massive du niveau d'éducation et d'emploi des femmes amorcé à partir des années 1960 à travers le monde. Dans cette dynamique, les niveaux de qualification des femmes ont augmenté plus vite et atteint des niveaux plus élevés que les hommes dans la majorité des pays de l'OCDE. La proportion de femmes de 25-34 ans diplômée du supérieur a très fortement augmenté par rapport à la génération des 45-54 ans, en général d'autant plus que la proportion de ces dernières est faible. Chez les hommes, la hausse est en moyenne nettement plus faible et inférieure à ceux des femmes. Les données Françaises sont proche de la moyenne des pays de l'OCDE avec environ 40 % de femmes entre 25 et 64 ans avec un diplôme du supérieur. Entre 2008 et 2018, la part des hommes de 25-34 ans en France avec un diplôme du supérieur est passée de 36 % à 43 % et celle de femmes est passée de 45 % à 51 % (OECD 2019a).

Dans les dernières données disponibles présentées par Claudia Goldin à la cérémonie pour son prix Nobel, la proportion de femmes ayant eu accès à l'université est supérieur à celle des hommes dans tous les pays de l'OCDE

(Figure 1).

Figure 1: College attendence by gender in 2022 in OECD country, from Goldin's Nobel Lecture (2023).



L'équilibre sur le marché du travail dépend de la demande pour les différents métiers et compétences de la population et donc de la capacité de l'économie à produire de bons emplois. Au départ de l'analyse de Goldin, il y a un modèle économique (offre/demande) dont il est bon de rappeler les bases: si la demande augmente face à une offre inélastique, le taux de salaire augmente (et réciproquement); si l'offre est élastique, l'effet sur le salaire est ambigu et dépend de la valeur de cette élasticité. La croissance économique mondiale des pays développés depuis les années 1960 s'est largement intensifiée en usage du capital, mais aussi en travail très qualifié et beaucoup moins en main d'œuvre moins qualifiée, dans une forme de "course" entre éducation et technologie modifiant l'équilibre général de l'économie (C. Goldin and Katz 2008; Acemoglu and Autor 2012; D. Autor, Goldin, and Katz 2020).

Notons que si la demande de travail qualifié n'augmente pas suffisamment pour absorber les travailleurs formés ou si les compétences ne correspondent pas à l'économie, le marché du travail peut se trouver dans une situation de "sur-diplôme" entraînant potentiellement du déclassement ou une compétition plus forte pour des emplois moins rémunérateurs, conduisant à l'éviction durable des moins diplômées des emplois qui leurs correspondent et/ou du chômage de masse (Groot and Maassen van den Brink 2000; Allen and van der Velden 2001; Büchel and Mertens 2004; McGuinness 2006; Pollmann-Schult 2005).

I.2 Intensification de la croissance en technologie et capital humain

Plusieurs forces économiques et institutionnelles alimentent ces dynamiques. Notamment l'intensification du commerce international accompagné d'une réduction importante des tarifs douaniers, de nouveaux systèmes monétaires internationaux[10] d'une libéralisation des mouvements de capitaux et des personnes[11]. L'automatisation des processus industriels et agricoles, les délocalisations, l'arrivée des technologies de l'information et des communications, le développement de la finance internationale, la création d'unions douanières, économiques et monétaires,… ont transformé les tissus économiques des différents pays en fonction de leurs avantages comparatifs. Des secteurs entiers ont disparu, d'autres sont apparus, modifiant la dynamique des territoires accentuant les inégalités, avec des conséquences plus ou moins lourdes pour celles et ceux qui ont le plus gagné ou perdu à ces changements.

[10] Des accords de Bretton-Woods en 1944 à la création de l'Euro, l'émergence de nouvelles puissances économiques et monétaires.

[11] Notamment au sein de l'Union Européenne. Cependant, les questions migratoires concentrent toujours - et probablement davantage encore dans la période récente - un niveau élevé de tensions entre intérêts économiques divergents et relations à l'identité nationale. Voir les travaux récents de Kovacic and Orso (2023) et Cavaillé and Van der Straeten (2023)

La hausse de demande pour des travailleurs qualifiés est telle que le rendement de l'enseignement supérieur s'est fortement accru - avec de fortes disparités entre les pays et au sein des pays - et ces diplômés et diplômées ont globalement pu trouver des débouchés sur le marché du travail, augmentant massivement le taux d'activité des femmes. Les revenus relatifs des diplômés de l'enseignement supérieur ont augmenté de manière spectaculaire dans les années 1980 et le différentiel entre les genres a diminué à la fois dans l'ensemble et pour tous les groupes d'âge et d'éducation dans les années 1980 et 1990 (Katz and Autor 1999).

Ces changements fondamentaux se poursuivent, nourissent et se nourissent d'autres transformations des sociétés dans une dynamique d'accroissement important des richesses et du niveau de vie général, mais également du nombre, de la proportion, et des profils de populations durablement éloignées du marché du travail.

La dégradation des conditions des personnes moins diplômées, ou aux compétences dépassées À l'opposé de la révolution du capital humain, la situation économique de celles et ceux moins diplômés, entrant sur le marché du travail ou travaillant dans des secteurs en déclin, s'est très fortement dégradée dans les pays développés, faisant naître un chômage de masse et structurel important. Ce fait social deviendra l'une des principales préoccupations des décideurs politiques conduisant à la transformations des États providence (Giuliano Bonoli 2010).

Les hommes étant moins diplômés que les femmes et les secteurs les plus exposés à la mondialisation étant surtout les secteurs manufacturiers et agricoles employant relativement plus d'hommes, leur situation économique s'est fortement dégradée depuis les années 1970. Quant aux femmes moins diplômées, elles constituent la majorité de l'accroissement des familles monoparentales (Stevenson and Wolfers 2007).

Dans une publication du *Journal of Economic Litterature*, Coile and Duggan (2019) recensent les données et travaux qui analysent la situation des hommes peu qualifiés *lorsqu'il n'y a plus d'emploi* et constatent qu'au-delà de la baisse des taux d'activité et des revenus, on observe un ralentissement de l'accroissement voire une diminution de l'espérance de vie, une augmentation de la mortalité précoce, des suicides, de l'usage de drogue, des incarcérations, un état de santé qui se dégrade et un recours plus important aux programmes d'invalidité.

On observe également une diminution du taux de mariage, en particulier parmi les hommes n'ayant pas fait d'études supérieures. De nombreux travaux documentent les effets de l'entrée sur le marché du travail lorsque l'économie est en récession (Maclean 2013; Oreopoulos, Von Wachter, and Heisz 2012; von Wachter 2020) et des effets durable d'une perte d'emploi (Lefranc 2003; Oreopoulos, Von Wachter, and Heisz 2012; Sullivan and Wachter 2009; Nafilyan 2016; Schmieder, Wachter, and Heining 2022). Les travaux montrent généralement des effets négatifs durables sur les trajectoires de vie des personnes (santé, emploi, mortalité), plus marqués parmi les populations défavorisées.

Binder and Bound (2019) argumentent que bien que les différences de demande de travail qualifié et non qualifié expliquent une partie des difficultés de ces hommes, elles ne suffisent pas à comprendre la dynamique générale de l'emploi et de l'éducation des hommes. En revanche, intégrer la dynamique des couples et des mariages dans cette dynamique apporte un éclairage supplémentaire. Dans une étude célèbre, Akerlof, Yellen, and Katz (1996) analysent l'augmentation des naissances hors mariage et s'interrogent sur les causes de la baisse des mariages précipités (*shotgun mariages*). Les auteurs considèrent également l'évolution des relations de pouvoirs entre hommes et femmes en convoquant un argument intéressant. Ils considèrent que les technologies contraceptives et abortives ont favorisé la révolution sexuelle où celles adoptant ces technologies pouvaient décider d'avoir des relations sexuelles sans devoir les conditionner à une promesse de mariage en cas de grossesse. Les relations sexuelles extra-maritales sont devenues de plus en plus fréquentes sans que les hommes n'aient besoin de promettre de s'engager pour en avoir. Formalisés dans une série de modèles théoriques, Selon Akerlof, Yellen, and Katz (1996), celles qui, pour diverses raisons, n'ont pas opté (ou pu opter) pour ces technologies ont subit une pression plus forte à des relations sexuelles hors mariage de la part d'hommes moins enclins à s'engager du fait de ce changement de pouvoir. Or ces femmes sont majoritairement moins éduquées et/ou racisées. Les modèles présentés sont conformes aux données et résultats présentés plus haut indiquant une hausse de la monoparentalité, une baisse des mariages précipités touchant particulièrement les femmes défavorisées.

En examinant les différences selon l'origine raciale, Coile and Duggan (2019) constatent des baisses plus importantes de la mortalité et de l'incarcération depuis 2000 parmi les populations noires et hispaniques. Ces groupes ont également enregistré des progrès éducatifs plus rapides. La situation des hommes gays s'est aussi particulièrement améliorée avec la réduction de l'incidence et la mortalité du VIH, de moindre discriminations dans l'éducation et le marché du travail réduisant les inégalités avec les hommes hétérosexuels (Badgett, Carpenter, and Sansone 2021). Dans une perspective queer et intersectionnelle, le groupe majoritaire des hommes blancs hétérosexuels souffre de son manque de réactivité face aux changements majeurs qui se sont opérés. C'est celui qui a le moins suivi la dynamique générale vers une augmentation du niveau de diplôme et subit le plus les conséquences des phénomènes que nous venons de décrire. À l'opposé, les hommes des groupes minoritaires (racisés et/ou non-hétérosexuels) ont accompagné le mouvement de la révolution discrète, améliorant leur situation et réduisant les différences moyennes avec le groupe qui constitue la "norme".

I.3 Restructuration des familles

Un point central de l'argument de Claudia Goldin (2006-May) est que l'augmentation du niveau d'éducation des femmes a pu se faire grâce à des chocs technologiques et institutionnels. Au dela de la contraception et de l'avortement discutés plus haut, les transformations du tissu économique ont également permis de réduire le temps nécessaire à la réalisation des tâches domestiques. L'aspect technologique ne s'arrête pas aux robots ménagers mais aussi à l'accès à des biens et services à moindre coûts - parfois subventionnés sur fonds publics - permettant d'externaliser certaines activités (garde d'enfant, ménage, soutien scolaire…) ou de consommer différemment (racheter des vêtements plutôt que coudre ou consommer des aliments préparés notamment) (Champagne, Pailhé, and Solaz 2015a; Greenwood, Seshadri, and Yorukoglu 2005; Bose, Jain, and Walker 2022).

Pour l'économie domestique, une part de ces services étaient (et sont encore) assurés dans l'économie informelle et leur marchandisation visait à les inclure dans l'économie formelle, mais aussi à créer des emplois pour les travailleurs (mais surtout travailleuses) peu qualifiées en subventionnant leur embauche. Enfin ils permettent de libérer du temps pour travailler aux personnes qui ont recourt à ces services, et notamment les familles plus aisées (Brück, Haisken-De New, and Zimmermann 2006; Cortés and Tessada 2011).

Pendant longtemps, la relation entre niveau d'étude des femmes et taux de mariage était décroissante (Lundberg, Pollak, and Stearns 2016). Elle s'est à présent retournée, dans plusieurs pays développés, lui donnant une allure de courbe en U (Bertrand et al. 2021). Il y a là un signe fort d'appariement assortatif et un rôle différent du mariage pour des familles de niveaux d'éducation différent. Les recherches en économie de la famille ont intégré des processus de décision plus complexes, intégrant des interactions stratégiques et un pouvoir de négociation, la possibilité de transfers d'utilité entre les membres, une complémentarité dans la fonction de production domestique et de capital humain des enfants (Lundberg and Pollak 1993; Pierre-André Chiappori 2017).

Empiriquement, l'assortativité des relations de couples se retrouve dans de nombreuses recherches (Mare 1991; Gonalons-Pons and Schwartz 2017) qui montrent par ailleurs d'autres fonctions latentes du mariage: un rôle d'assurance permettant de mutualiser les risques, notamment vis à vis d'une perte d'emploi (K. Huber and Winkler 2016; Clark, D'Ambrosio, and Lepinteur 2022), un puissant levier d'accumulation de capital via un accès facilité au marché des fonds prétables et des économies d'échelles dans la consommation et l'investissement immobilier (Bessière and Gollac 2020) ; autant de bénéfices qui, lorsqu'ils disparaissent, rendent les effets négatifs des ruptures plus importants et durables (Nieuwenhuis 2022; Bonnet, Montaignac, and Solaz 2024).

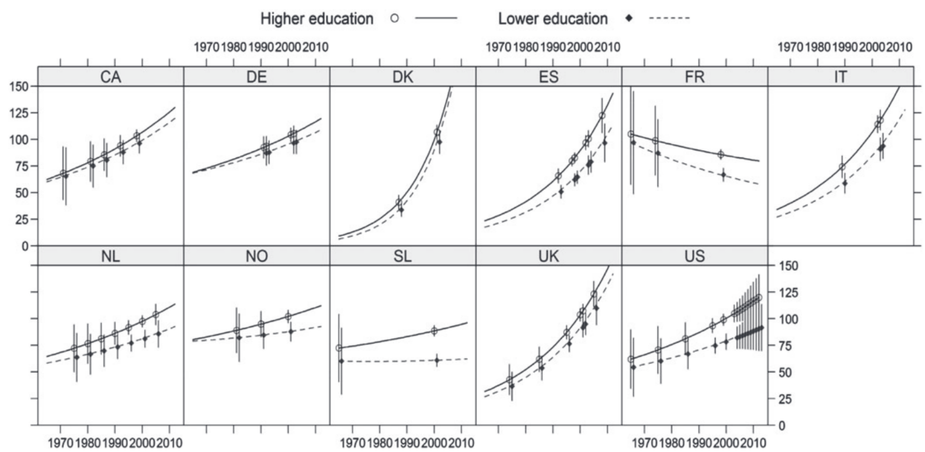
Usage du temps et investissement dans l'éducation des enfants La complémentarité entre partenaires de couple se retrouve dans les dynamiques d'investissement dans le temps passé à éduquer les enfants.

Dotti Sani and Treas (2016) exploitent des enquêtes internationales standardisées répétées menées dans différents pays de l'OCDE et comparent l'évolution du temps moyen passé à s'occuper des enfants entre pères et mères cohabitants d'enfants de moins de 13 ans de couples hétérosexuels suivant leur niveau d'éducation et contrôlant pour un certain nombre de caractéristiques. Nous reproduisons leur principaux résultats dans les figures ci-après extraites de cet article. On note tout d'abord que le temps moyen des femmes reste partout plus élevé que celui des hommes. À part en France, les mères passent en moyenne plus de temps à s'occuper de leurs enfants dans les données plus récentes et souvent d'autant plus qu'elles ont un niveau d'éducation élevé. Ce phénomène d'intense parentalité a des effets contrasté sur les inégalités et reflète des stratégies différentes (Verniers, Bonnot, and Assilaméhou-Kunz 2022). Les pères ont eux aussi généralement augmenté le temps passé à s'occuper des enfants mais l'écart entre niveau d'éducation est plus marqué que pour les femmes. Dans plusieurs pays, les groupes d'hommes moins diplômés n'ont pas du tout augmenté ce temps.

Figure 2: Evolution du temps moyen prédit à s'occuper des enfants, tiré de Dotti Sani and Treas (2016)

(a) Temps des mères

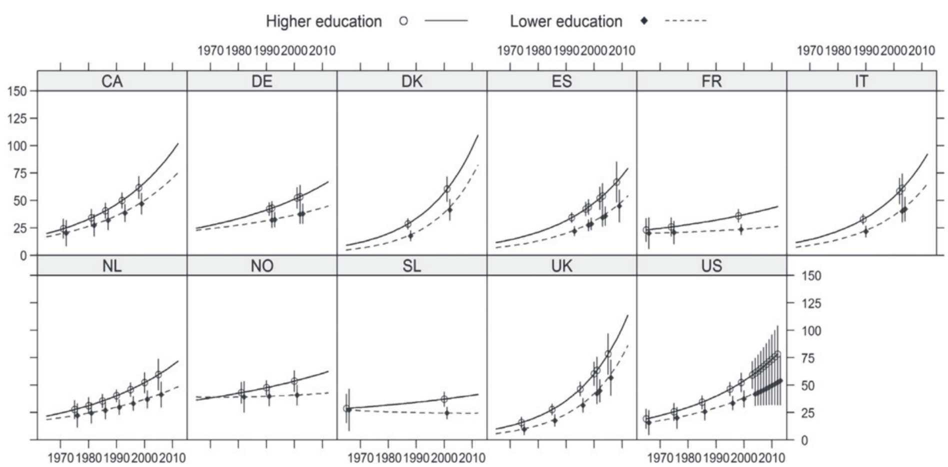
FIGURE 1. MOTHERS' PREDICTED CHILD-CARE MINUTES DAILY BY EDUCATION AND YEAR FOR 11 COUNTRIES.



Note. The predictions and confidence intervals are calculated from Model 3. The predicted values are adjusted by setting age, partnership status, number of children, age of youngest child, employment status, and day of week at the overall sample means.

(b) Temps des pères

FIGURE 2. FATHERS' PREDICTED CHILD-CARE MINUTES DAILY BY EDUCATION AND YEAR FOR 11 COUNTRIES.



Note. The predictions and confidence intervals are calculated from Model 3. The predicted values are adjusted by setting age, partnership status, number of children, age of youngest child, employment status, and day of week at the overall sample means.

Dans ces données, les couples français suivent clairement des tendances différentes de ceux d'autre pays et au global, les enfants de parents français passent de moins en moins de temps avec leurs parents bien que les pères éduqués compensent légèrement cette baisse. Champagne, Pailhé, and Solaz (2015a) analysent les données françaises entre 1985 et 2010 montre que sur ces 25 années, les femmes ont, en moyenne, consacré davantage de temps aux activités parentales, mais elles ont sensiblement réduit le temps dédié à l'entretien domestique. Les auteurs et autrices ajoutent que cette baisse tient surtout aux changements de leurs pratiques, et dans une bien moindre mesure à la progression de l'activité féminine et aux changements des structures familiales. La baisse du temps concerne la couture, le linge et la cuisine - trois tâches pouvant être remplacé par de la consommation - tandis que le temps de trajets a triplé. Les hommes se sont en moyenne davantage impliqués dans l'éducation des enfants, les pères peu ou non-participants devenant plus rares. Toutefois, la contribution des hommes aux autres tâches domestiques est demeurée stable. En 2010, les femmes effectuent ainsi la majorité des tâches ménagères et parentales – respectivement 71 % et 65 %. Cette inégale répartition montre des résistances à un partage plus égal des tâches dans la population. Dans un document de travail publié au *NBER*, Briselli and Gonzalez (2023) font l'hypothèse que les performances de genre de hommes et notamment leur refus de prendre une part plus importante à la production domestique sont à la fois un frein aux mariages, à la fertilité et à l'activité des femmes dans des analyses entre pays et sur données individuelles.

Réduction et persistence des inégalités de genre Plusieurs travaux permettent de mesurer à quel point la révolution discrète de Goldin a transformé le marché du travail en France également. En particulier, Meurs and Pora (2019) documentent l'évolution des inégalités de genre sur le marché du travail sur longue période. La Figure 3 tirée de cet article illustre parfaitement la tendance longue de réduction des inégalités. L'écart entre les femmes et les hommes s'est réduit tant sur le taux d'activité (panel a), le taux de chômage (panel b) que sur les taux de rémunérations, sur toute la distribution de salaires. En revanche, les inégalités sur le marché du travail s'accroissent massivement et durablement à l'arrivée d'un enfant. Ce phénomène de pénalité liée à l'enfant (*child penalty*) est un phénomène mondial (H. Kleven, Landais, and Leite-Mariante 2024), ne s'explique pas par l'effet propre de la maternité puisque cette pénalité existe autant avec des enfants adoptés. Elle n'existe en revanche pas parmi les couples d'hommes qui adoptent (H. Kleven, Landais, and Søgaard 2021). Parmi les couples lesbiens, le choc initial est divisé de moitié et partagé entre les co-parentes et se résorbe en 5 ans alors qu'il persiste pour les couples hétérosexuels (Andresen and Nix 2022). Les travaux de Lundborg, Plug, and Rasmussen (2017) mesurent également la pénalité liée à l'enfant mais comparent succès et échecs de procédures de procréation médicalement assistée à partir de données administratives exhaustives danoises. Ces analyses mesurent la même pénalité de genre que les travaux précédemment cités et documentent d'abord une baisse de participation au marché du travail et des heures travaillées puis une baisse du taux de salaire, entraînant une pénalité de naissance durable.

Les travaux de Meurs and Pora (2019) montrent que la pénalité liée à l'enfant est particulièrement prononcée pour les femmes du bas de la distribution de revenu alors qu'il se résorbe presque intègralement pour les femmes les plus aisées. Un résultat confirmé par Bazen, Xavier, and Périvier (2022) en comparant les groupes d'éducation parmi les entrants sur le marché du travail en France.

La demande de travail qualifié a fortement augmenté et encouragé les générations à s'éduquer pour y répondre mais les femmes ont beaucoup plus réagit à cette hausse de demande (Voir ci-après). Cependant, si le nombre de femmes entrant dans l'enseignement supérieur a augmenté la part de femmes parmi les cadres, les postes de direction, etc. il n'a pas éliminé les inégalités et les discriminations en raison du genre sur le marché du travail (Darity and Mason 1998 ; Juhn and McCue 2017).

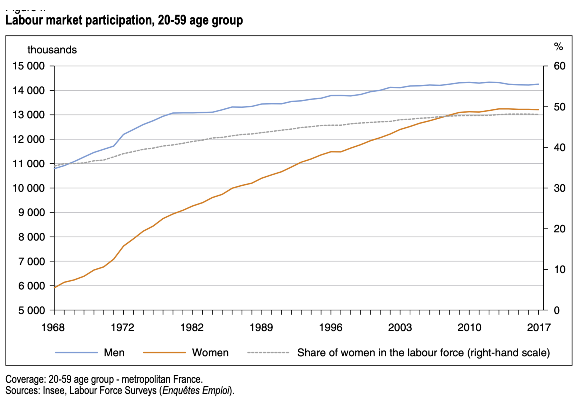
Lorsqu'on compare les hommes et les femmes issues des mêmes formations prestigieuses, les écarts de rémunérations entre hommes et femmes sont plus élevés et s'acroissent plus rapidement (Bertrand, Goldin, and Katz 2010). En France, Gobillon, Meurs, and Roux (2015) analysent les différences d'accès aux emplois par genre et salaires potentiels et montre que les femmes ont moins de chance d'accéder à des emplois du haut de la distribution de revenu.

Des inégalités persistent à tous les niveaux et de nombreux travaux documentent un accroissement des inégalités de réussite scolaire, d'accès à l'enseignement supérieur, en lien avec l'origine sociale, le genre, les inégalités territoriales et l'origine raciale (Corak 2013; Cnesco 2016). Les inégalités de revenus et de patrimoine se sont également fortement accrues aux États-Unis et dans d'autres pays développés, mais moins en France (au moins jusqu'aux dernières années) (Hoffmann, Lee, and Lemieux 2020 ; Chancel and Piketty 2021). Cependant, la reproduction sociale, c'est-à-dire le lien entre les inégalités "en coupe" et la mobilité intergénérationnelle s'est fortement accrue ; un phénomène surnommé "*The great Gatzby curve*"[12]. France stratégie a récemment publié une étude de Dherbecourt and Flamand (2023) sur les inégalités en France et montre dans des analyses en décomposition que l'origine sociale est, de loin, celle qui a le plus d'effet sur les écarts de revenus à l'âge adulte. Le genre arrive toujours en second malgré une réduction tendancielle des écarts de revenus.

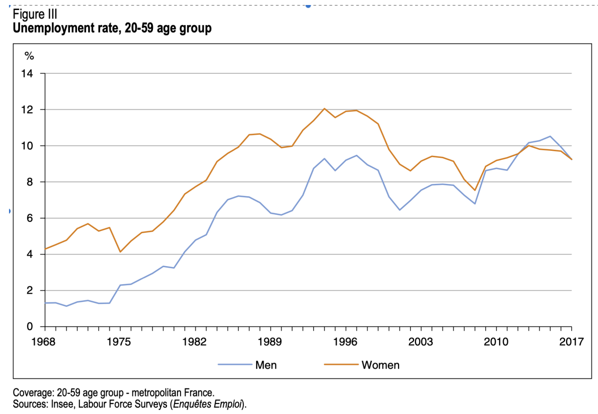
[12] Dans une présentation d'Alan Krueger, alors conseillé économique du président Obama. Voir Durlauf, Kourtellos, and Tan (2022)

Figure 3: La réduction des inégalités de genre sur le marché du travail et la pénalité liée à l'enfant

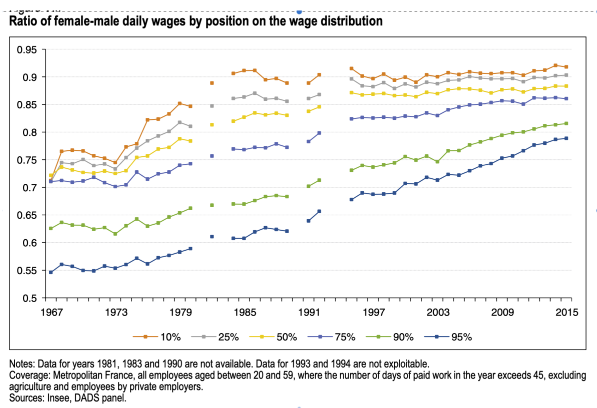
(a) Hausse massive de la participation au marché du travail



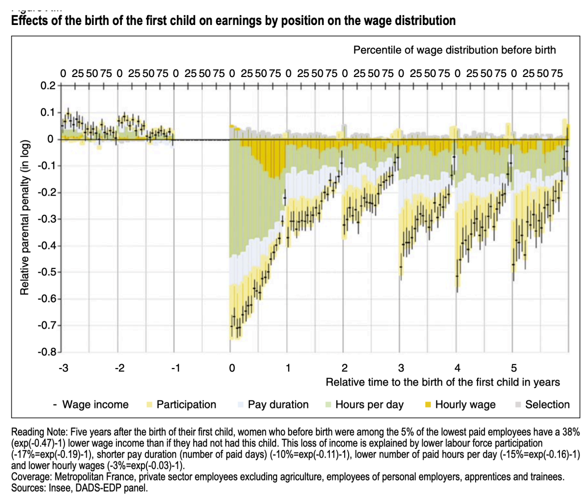
(b) Disparition de l'écart de taux de chômage



(c) Réduction des inégalités des taux de salaire



(d) pénalité liée à l'enfant par niveau de revenu



L'émèrgence des familles monoparentales Au cours des cinquante dernières années, la part de ces ménages a plus que doublé dans de nombreux pays de l'OCDE et représente maintenant environ un quart des foyers avec enfant dans l'Union Européenne (Figure 4. En France, elle est passée de 9 % en 1974 à 24,7 % en 2020, soit 2 millions de ménages, dont 82 % sont des femmes seules avec leurs enfants (INSEE 2021a). Il y a trois situations possibles menant à la monoparentalité dont la répartition a changé avec les transformations de la société : le veuvage en est de moins en moins souvent à l'origine, alors que les divorces ou les séparations de couples concubins sont de plus en plus fréquents. En France[13], 78 % proviennent d'une séparation, 16 % d'enfants nés en dehors d'une relation et 6 % du décès d'un des parents. Il est à noter que certaines séparations peuvent être liées à l'incarcération d'un parent (Désesquelles and Kensey 2006), ou le renvoie dans le pays d'origine pour les parents sans papiers (Carayon 2018 ; Shutes 2022).

Ainsi, les profils des familles monoparentales se diversifient : parents divorcés ayant la garde des enfants, jeunes mères n'ayant jamais vécu en couple, veuves ou veufs dont les enfants sont relativement grands, ou encore couples non-cohabitants, etc. Avant d'atteindre l'âge d'un an, 8 % des enfants nés en 2011 ne vivent pas avec leurs deux parents (Pailhé, Panico, and Heers 2020). Ces modes d'entrée dans la monoparentalité se traduisent par des perceptions très contrastées de la vie familiale antérieure et des perspectives également très différentes pour les familles concernées (Baronnet et al. 2021). Bien que la plupart des enfants de familles monoparentales soient élevés par

[13] Costemalle (2017) d'après des données INSEE de l'enquête Familles et Logements 2011.

leurs mères, le nombre de familles monoparentales gérées par des pères augmente, mais reste inférieur à 5 % des ménages, et leur situation économique est souvent bien meilleure que celle des mamans-solos[14].

Derrière ce phénomène, il y a notamment des écarts importants de niveaux d'éducation car la proportion de familles monoparentales s'est surtout accrue parmi les femmes les moins diplômées (Acs, Lhommeau, and Raynaud 2015). En 1990, la proportion de "mères isolées" était la même pour celles ayant un diplôme Bac +3 et celles ayant au plus le brevet des collèges, soit 10 %. En 2010, ces proportions sont passées à 11 % pour les premières et 24 % pour les secondes alors que la proportion de femmes diplômées dans la population a massivement augmenté. Entre 1999 et 2019, la proportion de femmes de 25-54 ans diplômées du supérieur est passée de 24,3 % à 46,3 %, leur taux d'activité est passé de 81,2 % à 88,6 % mais le taux d'emploi de celles n'ayant pas un niveau BAC est passé de 57,7 % à 52,6%[15].

La participation au marché du travail des femmes mariées et en concubinage a fortement augmenté, rejoignant les taux d'activité des familles monoparentales historiquement plus élevés en France. On retrouve alors des différences fortes de niveau de vie, de revenus, d'emploi entre les femmes mariées et en couples et les mères isolées qui confondent en partie les effets de ces différences de profils et ceux de la monoparentalité. Lorsque sont prises en compte l'ensemble de ces caractéristiques des mères, à savoir leur âge, le nombre et l'âge de leurs enfants, leur niveau de diplôme, leur statut matrimonial, la taille de l'unité urbaine et le taux de chômage de leur département de résidence, les mères isolées ont toujours une probabilité plus forte d'être actives que les mères en couple, mais une probabilité plus faible d'occuper un emploi (Acs, Lhommeau, and Raynaud 2015).

I.4 Les réformes des États providence

Des années 1950 au milieu des années 1970, des prestations relativement généreuses et peu conditionnées ont pu être accordées aux familles monoparentales, alors perçues comme "méritantes" Bullock, Williams, and Limbert (2003). À cette période, les familles monoparentales, pour beaucoup veuves de guerre, restaient relativement rares, dans une société marqué par le modèle de "Monsieur Gagnepain" (*Bread-Earner model*) où le niveau d'éducation formelle et la participation des femmes au marché du travail restaient limités (Claudia Goldin 2006-May; Périvier 2022b). Aux Etats-Unis, les années 1960 sont marquées par le développement d'un grand nombre de programme fédéraux, initiés sous la présidence de J.F. Kennedy et poursuivis par le président Johnson dont le leitmotiv était de faire la "Guerre contre la pauvreté" (*War on poverty*). L'allocation familiale américaine[16] mise en place en 1936 lors du "*New Deal*" se voit alors complétée en quelques années par un programme de bons alimentaires (SNAP), d'assurance santé (Medicare, Medicaid), de repas gratuit à l'école pour les enfants et d'un supplément au revenu d'activité (EITC) (Robert A. Moffitt 2015). Cependant, la notion de "pauvres méritants" (*deserving poor*) détermine déjà fortement les aides versées et se materialise notamment par des aides plus généreuse aux veuves qu'aux mères non mariées. Au Royaume-Uni, l'État providence de l'immédiat après-guerre a introduit des prestations nationales pour les veuves et les mères veuves, qui assuraient un soutien après le décès du mari, mais qui pouvaient être complétées par des revenus sans perte de pension. Les autres familles monoparentales dépendaient des prestations d'aide sociale, qui étaient moins élevées, soumises à des conditions de ressources et versées jusqu'à ce que le plus jeune enfant ait 16 ans, ou 18 s'il suivait encore un enseignement à temps plein (Knijn, Martin, and Millar 2007).

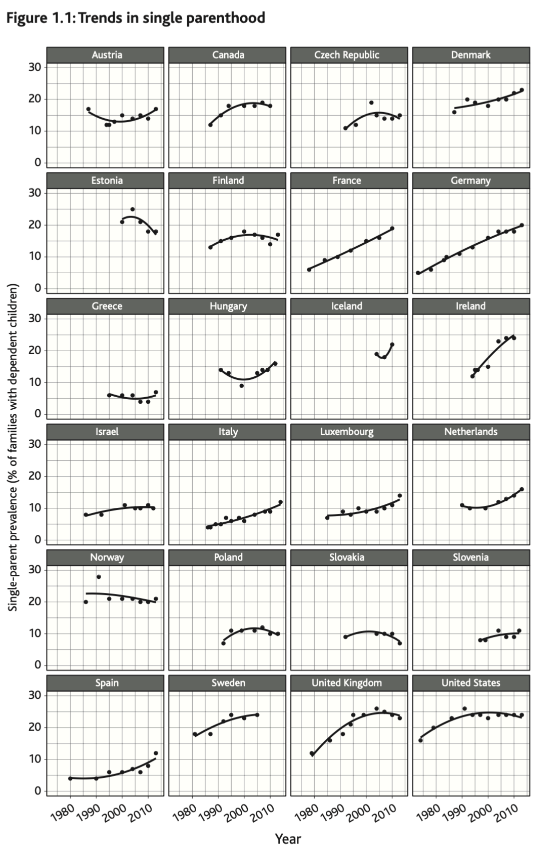
[14] Par exemple, les pères célibataires sont plus susceptibles de posséder leur maison (50 % contre 25 %), d'être en emploi (81 % contre 67 %), et lorsqu'ils le sont, de travailler en tant que cadres (18 % contre 10 %) (DREES 2021)

[15] Source : Insee, recensements harmonisés de la population 1999 et 2019 (exploitations complémentaires) et recensement 2017 de Mayotte, au lieu de résidence. Champ : Population des ménages âgée de 25 à 54 ans, hors élèves et étudiants.

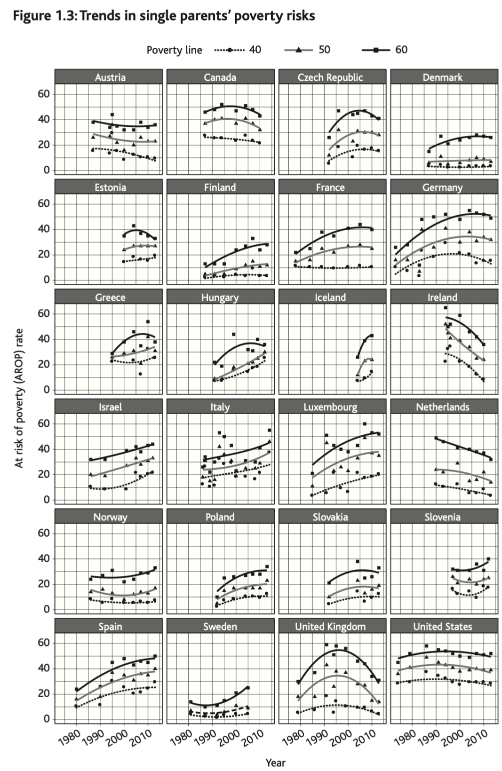
[16] AFDC - *Aid for families with dependent children*.

Figure 4: De plus en plus de familles monoparentales en situation de vulnérabilité, d'après Nieuwenhuis and Maldonado (2018)

(a) Hausse linéaire de la part de familles monoparentales



(b) Augmentation de la pauvreté monétaire des familles monoparentales



Le tournant de l'activation Depuis le début des années 1990, les pays de l'OCDE ont mis en place des politiques "d'activation" des dépenses sociales afin de (re)mettre une partie de la population inactive en emploi et favoriser leur autonomie économique (Giuliano Bonoli 2010). L'émergence de ces politiques n'est pas sans lien avec le développement d'une logique de réduction et rationalisation de la dépense publique motivées par des problèmes d'endettement public et l'émergence de la nouvelle gestion publique (New Public Management) qui s'est imposée dans les pays anglophones et scandinaves (White 2019). Il en découle un renforcement et une institutionnalisation de la logique d'évaluation de la mise en œuvre et des résultats des politiques (Lacouette Fougère and Lascoumes 2013; Bozio 2014; P.-H. Bono et al. 2021).

Les néolibéraux et conservateurs argumentaient que les politiques "passives" de protection sociale pouvaient décourager la recherche d'emploi et freiner la croissance économique en maintenant les individus dans une situation de dépendance vis-à-vis de l'État et en détournent des ressources du secteur productif (Midgley 1999). Les politiques actives du marché du travail ont été développées dans le but de promouvoir une approche plus individualiste et responsabilisante. L'activation s'appuie sur l'idée que, sur le marché du travail, la rencontre entre l'offre et la demande d'emploi peut être entravée notamment par une inadéquation entre les compétences recherchées par les entreprises et les profils des candidats et candidates ou une information imparfaite, telle que les demandeurs d'emploi n'ont pas connaissance de toutes leurs possibilités ni des méthodes pour effectuer une recherche efficace (Cahuc, Carcillo, and Zylberberg 2014). Mais les principes d'activation sont aussi motivés par la question de l'aléa moral : les demandeurs d'emploi indemnisés, ou toute autre personne recevant une prestation sociale, pourraient réduire leurs efforts de recherche tant que ce revenu de remplacement est disponible. Cela conduit à introduire des contrôles, éventuellement assortis de sanctions, pour inciter les personnes à maintenir leur activité de recherche à un niveau élevé. Parfois, les politiques d'activation penchent vers l'une ou l'autre conception. Mais souvent, les deux approches sont jointes : elles outillent et soutiennent tout en contrôlant.

L'activation touche un large éventail de politiques, de la fiscalité des ménages et des entreprises à la solidarité, aux politiques familiales en passant par l'éducation, la formation professionnelle, etc. Elle modifie progressivement et structurellement les États providence d'une logique d'assurance des risques sociaux (*welfare*) à une conditionnalité plus forte à la participation à la société, en particulier à travers le travail. Le *workfare* désigne la combinaison des programmes d'aide sociale avec des obligations de travail pour les bénéficiaires. Le *workfare* se distingue ainsi du *welfare*, qui est basé sur l'octroi de prestations sociales sans contrepartie obligatoire. En adoptant le *workfare*, les gouvernements cherchent à promouvoir l'autonomie des bénéficiaires, à favoriser leur réinsertion professionnelle et à réduire la dépendance aux prestations sociales. Le *workfare* englobe en particulier les politiques de *welfare-to-work* qui désignent les programmes d'accompagnement à la reprise d'activité pour les bénéficiaires de minima-sociaux.

Ces réformes ont généralement procédé par fusion, suppression, substitution d'aides existantes. Les nouveaux systèmes sont également moins généreux pour les personnes sans activité et les dispositifs d'aides sont souvent bornés en durée et davantage conditionnés à un suivi ou des obligations de recherche d'emploi, formations, etc. (Chappell 2010 ; Eydoux and Letablier 2009 ; Ziliak 2015).

Ces programmes sont également souvent universels i.e. ils concernent toutes les personnes en âge et capacité de travailler et s'accompagnent parfois d'autres aides ciblées (e.g. en France l'allocation parent isolé transformée en RSA majoré) ou dispositifs complémentaires dédiées à certaines populations (typiquement, les jeunes ni en emploi, ni en formation, ni en éducation, les bénéficiaires de minima-sociaux, les familles monoparentales, les personnes en situation de handicap, les chômeurs longue durée ou séniors,…). Knijn, Martin, and Millar (2007) documentent dans plusieurs pays la suppression de prestations de sécurité sociale qui étaient spécifiquement destinées aux familles monoparentales, ou leur fusion avec les prestations destinées à protéger contre le chômage. Les nouveaux systèmes sont généralement moins généreux, bornées en durée et davantage conditionnés à un suivi ou des obligations de recherche d'emploi, formations, etc. Cette "recatégorisation des risques", qui a donc également touché d'autres groupes auparavant largement exclus d'obligations de recherche d'emploi[17] (Dwyer et al. 2020), a officiellement mis fin au statut spécial dont bénéficiaient jusqu'alors les mamans-solos - ou, plus généralement, les mères sans partenaire soutien de famille - dans les systèmes de sécurité sociale de nombreux pays, à des degrés divers (Jaehrling, Kalina, and Mesaros 2015).

Parmi les exemples de réformes les plus importantes de la première vague, on trouve le *New Deal* britannique (Gregg, Harkness, and Smith 2009), les réformes *Personal Responsibility and Work Opportunity Act* (PRWOA) signées sous l'administration Clinton (Peterson 1997 ; Robert A. Moffitt 2003 ; Carcasson 2006), les réformes Hartz en Allemagne (van Gestel and Herbillon 2007 ; Vail 2008 ; Engbom, Detragiache, and Raei 2015), la privatisation des services d'accompagnement en 1998 avec la création du « Job Network » en Australie (de Gendre, Schurer, and Zhang 2022).

Les plus grosses réformes de ce type sont adoptées en France à partir de 2008 avec d'une part la création du RSA et d'autre part la création de Pôle Emploi en 2008, à partir de la fusion de l'ANPE et de l'UNEDIC. Auparavant, l'activation en France a d'abord et surtout été pensées dans une politique d'offre visant à réduire les coûts du travail notamment pour les bas salaires, augmenter le taux d'emploi par le partage du temps de travail (introduction de temps partiels et contrats courts, 35h) (Kramarz, Nimier-David, and Delemotte 2022). La question des incitations est devenue plus centrale à partir de la fin des années 1990 cristallisé autour de la question des "trappes à inactivités" et l'influence du "*making work pay*" des pays anglophones, objet du Chapitre 3.

[17] e.g. en Angleterre le programme *Pathway-to-Work* pour favoriser l'emploi des personnes en situation de handicap (Adam et al. 2008)

II Investissement social et la lutte contre la pauvreté en France

Un premier aspect distinctif de l'investissement social réside dans sa terminologie même : le remplacement du terme "dépense" par "investissement" clarifie la sémantique : les dépenses sociales produisent des effets dans la société et ce sont ces retours qui justifient l'intervention de l'État. Cette approche vise à réconcilier l'économique et le social, souvent opposés par les acteurs politiques, en mettant l'accent sur la capacité de la protection sociale à stimuler l'activité et l'emploi, contribuant ainsi à un financement durable des dépenses sociales (Palier 2014). Les dépenses de protection sociale peuvent alors être catégorisées selon leur dimension productive (De Deken 2016; Vandenbroucke and Vleminckx 2011a), ou dans leur rôle d'accumulation et de préservation du capital humain des individus (A. C. Hemerijck 2014).

Les familles monoparentales occupent une place particulièrement centrale dans la littérature sur l'investissement social, car elles sont souvent la cible des politiques visant à concilier vie familiale et professionnelle. Les modes d'accueil formels, notamment, permettent à la fois de libérer les parents pour travailler et de favoriser le bon développement des enfants grâce à une prise en charge de qualité. L'objectif est également de réduire les inégalités sociales dès le plus jeune âge (Van Lancker 2013).

À travers toute l'Europe, d'importants budgets sont alloués à l'activation des familles monoparentales ainsi qu'à la garde d'enfants, aux congés parentaux et autres types de congés, et à la promotion de l'égalité des chances dans l'enseignement et la santé (Cantillon and Van Lancker 2011). La dimension intergénérationnelle occupe généralement une place centrale dans l'investissement social, car la pauvreté et le chômage des parents peuvent se transmettre aux enfants par différents canaux (Ermisch, Francesconi, and Pevalin 2004; Hartley, Lamarche, and Ziliak 2022).

L'investissement social dans la petite enfance La France se démarque des autres pays de l'OCDE par des taux historiquement élevés de fécondité et de participation des femmes sur le marché du travail, ainsi que par des politiques familiales et de garde relativement généreuses (OECD 2019b). Le système français offre aux parents un large éventail d'options de garde et de mesures de soutien aux jeunes parents. Parmi celles-ci figurent le congé parental partagé (PrePare), des crédits d'impôt pour les dépenses de garde d'enfants, la Prestation d'accueil du jeune enfant (PAJE) et les allocations familiales à partir du deuxième enfant. Les services de garde formels sont accessibles aux enfants dès l'âge de trois mois, et l'école maternelle est désormais obligatoire à partir de trois ans. En 2020, le gouvernement français a investi 6,6 milliards d'euros dans les crèches collectives, 4,6 milliards d'euros dans les assistantes maternelles et 1,7 milliard d'euros dans les crédits d'impôt (Ishii et al. 2023). Si les modes d'accueil étaient pleinement utilisés, les crèches pourraient accueillir 471 000 enfants (soit 20 % des enfants de moins de trois ans) et les assistantes maternelles 744 000 enfants (soit 33 %). En 2020, le taux de couverture global pour la garde d'enfants de moins de trois ans était d'environ 60 %, et les dépenses publiques directes totales pour la garde d'enfants en 2020 s'élevaient à 14,3 milliards d'euros.

Des enquêtes représentatives révèlent que les parents préfèrent les crèches aux autres modes de garde, mais leur accès est fortement contingenté et inégal entre les territoires et les catégories de population (Bouteillec, Kandil, and Solaz 2014a; Laporte, Crépin, and Hilairet 2019; Cartier et al. 2017). Depuis le début de nos travaux jusqu'à aujourd'hui, les gouvernements successifs ont déployé des efforts considérables pour augmenter l'offre et améliorer l'accès aux modes de garde formels:

- La convention d'objectif et de gestion 2018-2022 de la Branche famille affichait un objectif ambitieux d'augmentation de l'offre ;
- En 2017, l'Association des Maires de France s'est vue confier une mission sur la transparence dans l'attribution des places en crèche (Laithier 2018) ;
- À l'automne 2018, le président français Emmanuel Macron a présenté une stratégie de prévention et de lutte contre la pauvreté qui s'est fixée une ambition particulière[18] : éradiquer la pauvreté grâce à une stratégie d'investissement social centrée sur l'enfant, visant à favoriser les investissements en capital humain pour les familles défavorisées[19].

[18] Voir la présentation de la stratégie pauvreté sur le site https://solidarites-sante.gouv.fr

[19] Dans le plan initial, les interventions pour les adultes en situation de pauvreté étaient principalement axées sur l'augmentation de l'accès à la formation professionnelle et au soutien social.

- Le cycle de séminaire “Premiers Pas” porté par la Cnaf, France Stratégie et le Haut Conseil de la Famille, de l’Enfance et de l’Age (HCFEA) et les travaux de la commission “1000 jours” sous la direction de Boris Cyrulnik ont permis de dresser un état des lieux de la petite enfance et dessiné les contours d’un service public de la petite enfance en cours de déploiement.
- Le 1er juin 2023, la Première ministre a annoncé un nouveau plan visant à créer 200 000 places en crèche d’ici 2030[20].

L’investissement social et lutte contre la pauvreté des adultes Dans la perspective de l’investissement social, la lutte contre la pauvreté est au centre des préoccupations pour au moins trois raisons :

- **Son incidence** : la pauvreté affecte en particulier les enfants et jeunes adultes d’une part, et les familles monoparentales d’autre part (Ermisch, Francesconi, and Pevalin 2004).
- **Ses effets sur le capital humain de la population** : la pauvreté détruit et empêche l’accumulation de capital humain des populations affectées (Schilbach, Schofield, and Mullainathan 2016b). Elle entraîne également une stigmatisation, des difficultés de santé physique et mentale, ainsi qu’une moindre planification financière à long terme (Campbell et al. 2016a).
- **Son rôle d’objectif des politiques publiques menées**: la lutte contre la pauvreté est une priorité des politiques publiques en raison de ses implications profondes sur le bien-être social et économique de la population.

Ainsi, parce que la pauvreté touche des publics fragiles particulièrement exposés aux risques d’exclusion, les privant durablement du capital humain et social nécessaire à leur participation à la société, la lutte contre la pauvreté est un objectif premier de l’investissement social.

Pour y parvenir, l’investissement social intervient, d’une part, en amont (petite enfance, éducation, soutien à la fonction parentale, etc.) pour permettre aux jeunes de se doter de compétences leur permettant à l’avenir d’être moins exposés et plus résilients face aux risques sociaux (Hartley, Lamarche, and Ziliak 2022). D’autre part, il s’engage dans l’accompagnement et la formation des adultes en situations de pauvreté, de chômage et d’exclusion afin de faciliter leur réinsertion par la participation au marché du travail (Jenson 2010; Giuliano Bonoli 2011).

L’investissement partage avec les politiques d’activation le principe selon lequel le moyen le plus efficace de sortir durablement de la pauvreté consiste à permettre aux individus de participer au marché du travail (Esping-Andersen 1990). Mais il s’en distingue en insistant sur l’idée que la participation sociale ne peut aller sans protection des plus vulnérables. Les politiques d’activation doivent donc être interrogées au regard de leur impact sur le retour à l’emploi, mais aussi sur la sortie de la pauvreté, le revenu et les conditions de vie des populations affectées (Avenel et al. 2017; Garritzmann, Häusermann, and Palier 2022).

Enfin, l’investissement social intègre une composante de protection ou maintien du niveau de capital humain à travers des transferts monétaires assurantiels où sous forme de minima-sociaux. Cette composante a fait l’objet de nombreux débats académiques car elle relève typiquement des interventions classiques des États providences tels que définis par Esping-Andersen (1990). Dans le cadre de l’investissement social, Hemerijck@2014 parle de fonction d’amortisseur (*Buffer*), et ce rôle est fondamental.

La protection des revenus et l’investissement l’investissement social sont deux dimensions qui, au moins potentiellement, divisent les acteurs de manière distincte et génèrent ainsi des possibilités d’échange politique et d’accords ambigus. En conséquence, la modernisation du modèle conservateur de “Monsieur Gagnepain” ne suit pas une trajectoire linéaire: elle peut aller dans la direction d’au moins deux différents de développement politique : le *familialisme optionnel* ou le *défamilialisme*, selon la façon dont les politiques d’investissement social sont combinées avec les politiques de remplacement du revenu (Kuitto 2016; Häusermann 2018).

[20] Le monde, 2023-06-01

La question centrale de l'évaluation des politiques publiques G. Bonoli (2012) rappelle que la notion d'investissement social est largement basée sur la promesse que l'argent dépensé aujourd'hui rapportera à un moment dans le futur. La crédibilité de cette démarche dépend de l'existence de travaux empiriques capables de démontrer ces retours.

La France se distingue dans sa culture de l'évaluation par un fort attachement à l'analyse de processus et de résultats pour lesquels on retrouve fréquemment des outils qualitatifs orientés par des préceptes de pertinence, de cohérence et d'efficacité, ainsi que des méthodes d'observations participatives (Lacouette Fougère and Lascoumes 2013; Desplatz and Lacouette Fougère 2019). Il existe en France une vision répandue selon laquelle l'évaluation ne contribue pas utilement à la réforme, et le fait que les évaluations semblent avoir peu d'impact sur la réforme de leur objet ne semble déranger personne (Delahais and Lacouette-Fougère 2019).

L'investissement social fait la promesse que la dépense sociale engagée aujourd'hui rapportera à un moment donné dans le futur si elle est orientée vers la dotation des individus en capital humain. Au delà du *marketing politique* – qui tend à vider le concept de sa substance tant celui-ci se retrouve parfois dévoyé (Damon 2015; Nicole-Drancourt 2015) – il me semble qu'il ne peut avoir de traduction opérationnelle qu'à la condition – forte – d'être capable de démontrer ces retours. La notion de rendement de l'investissement social reste floue même si ce concept aspire à une acception monétaire, une analyse de ses coûts et bénéfices où seraient valorisés l'ensemble des effets – y compris non marchands – de l'investissement social. L'approche évaluative à la française a pu se trouver en défaut face à un tel challenge. La description de l'évolution des situations depuis la mise en œuvre des dispositifs ne constitue pas une mesure de l'impact et le constat que des objectifs ont été remplis ne garantit pas que ce soit le résultat de la politique mise en œuvre. Pas plus que la comparaison entre les bénéficiaires et non-bénéficiaires d'une politique publique n'apporte une bonne mesure de l'effet de cette dernière, sauf à y intégrer précisément des conditions d'évaluation appropriées.

Dans les travaux menés à France Stratégie, j'ai souligné qu'il était nécessaire de démontrer les retours sur investissement pour que l'investissement social gagne en crédibilité et devienne un outil plus opérationnel (Heim 2017). Cependant, une telle démonstration nécessite un changement radical d'approche par rapport aux travaux empiriques disponibles jusqu'à présent.

Définir les retours sur investissements sociaux intègre deux chantiers de nature assez différente. Le premier consiste à estimer les effets, directs et indirects, des mesures avec des méthodes d'estimations adaptées permettant d'interpréter ces résultats comme un lien de cause à effet. Le second consiste à donner une valeur monétaire à ces effets, à les actualiser puis les sommer et enfin, les comparer aux coûts différentiels des mesures, c'est-à-dire réaliser une évaluation coûts-bénéfices.

Denis Fougère et moi avons mené un groupe de travail et auditionné une centaine d'experts et expertes pour définir ce qui pourrait constituer un guide pour mesurer le rendement de l'investissement social. Nos conclusions peuvent être résumées de la façon suivante (Fougère and Heim 2019):

Tout d'abord, il n'y a pas d'obstacle structurel à appliquer les méthodes de calcul socioéconomique à l'investissement social : tout comme un projet d'infrastructure, un investissement social présente la caractéristique de générer des bénéfices sur un horizon éloigné — et pour certains non marchands — qu'il s'agit d'actualiser et de monétiser. Ensuite, l'application du calcul socioéconomique aux investissements sociaux se heurte aujourd'hui à la difficulté d'estimer les effets bruts de ces politiques, avant même actualisation et monétisation. Souvent diffus et hétérogènes au sein des populations, ces effets sont aussi plus compliqués à anticiper que les impacts d'une nouvelle infrastructure de transport, par exemple. Par conséquent, nous argumentions qu'il faudrait d'abord mettre en place les infrastructures de données permettant des suivis adaptés et développer à grande échelle les évaluations d'impacts et expérimentations aléatoires de terrain, en particulier pour les politiques publiques à forts enjeux. Alors seulement pourrions-nous alimenter des modèles plus structurés adossés à des valeurs tutélaires, à l'instar de ce qui se fait au Washington State Institute notamment.

La Cnaf, France Stratégie et le Liepp (Sciences Po Paris) ont collaboré sur la notion d'investissement social, visant notamment à répondre aux critiques telles que celle exprimée par Damon (2015), soulignant la dimension de "marketing social" (*social washing*) et à construire autour de cette notion une stratégie plus opérationnelle pour l'action publique (Avenel et al. 2017). Cela s'est concrétisé notamment par la promotion et le développement d'évaluations scientifiques des effets des politiques mises en œuvre.

Bien que les expériences aléatoires de terrain représentent une part faible, mais croissante, des évaluations réalisées en France (P.-H. Bono et al. 2021), la Cnaf, tout en étant familière avec ces méthodes, s'est davantage engagée dans le soutien de l'approche expérimentale, notamment via le soutien important apporté à cette thèse et aux projets menés. Ces derniers sont eux-mêmes désormais affirmés comme constituant un investissement social, comme le souligne la Convention d'objectif et de gestion entre la Branche famille et l'État pour la période 2023-2027. Cette convention inscrit notamment la possibilité pour les territoires de solliciter la Cnaf pour le recours aux algorithmes d'affectation des places en crèches développés dans le cadre de cette recherche[21], comme décrit dans le chapitre 1.

III Problématique générale

En 2018, au début de cette thèse, la population française compte 63,14 millions d'habitants, parmi lesquels 25,25 millions ont un niveau de vie inférieur à 18 989 euros annuels, soit 1 582 euros par mois, selon les données de l'Enquête Revenus fiscaux et sociaux de l'Insee (DREES 2021). Sur ce total, 9,3 millions de personnes, représentant 14.7% de la population totale, vivent en situation de pauvreté monétaire, définie par un revenu par unité de consommation inférieur à 60 % du niveau de vie médian, soit 1 063 euros par mois en 2018.

Alors que l'investissement social occupe une place de choix dans la conception des politiques publiques dans la plupart des pays de l'OCDE et revêt une importance centrale dans d'autres domaines des sciences sociales, cette notion reste relativement marginale en économie. Néanmoins, dans cette thèse, je l'adopte comme cadre général pour concevoir et évaluer les politiques de la petite enfance ainsi que les politiques actives du marché du travail destinées aux familles monoparentales. Dans les deux cas, l'investissement social fournit des objectifs politiques ainsi que des critères normatifs permettant de juger les effets souhaités, les externalités et d'évaluer le succès ou l'échec des politiques mises en œuvre. Pour ce faire, je concentre mon analyse sur les trois principaux récits du paradigme de l'investissement social :

1) Investir dans la petite enfance pour réduire les inégalités et rompre le cycle intergénérationnel de la pauvreté ;
2) L'activation comme moyen de favoriser l'emploi et de réduire la pauvreté ;
3) Les suppléments de revenu pour rendre le travail rémunérateur.

En France, la littérature sur les effets des investissements sociaux dans la petite enfance ou dans les politiques actives du marché du travail est relativement limitée (P.-H. Bono et al. 2021). En particulier, il n'existe pas de preuves expérimentales des effets de l'accès à la garde d'enfants ou des programmes de transition du bien-être au travail pour les familles monoparentales. Cette thèse vise à combler cette lacune en proposant une approche expérimentale pour les deux cas. Ma principale contribution réside dans l'utilisation de méthodes économiques de pointe pour répondre à la question de recherche suivante :

Comment les politiques d'investissement social affectent-elles la dynamique de la pauvreté en France ?

Chaque chapitre aborde un puzzle spécifique :

- **Puzzle n°1** : Pour que les politiques petite enfance puissent atteindre les objectifs qui leurs sont assignés, il est nécessaire de s'assurer que ceux qui peuvent en bénéficier le plus y ont accès. Dans le cas des crèches, la rareté impose de faire des choix, et chaque choix implique un coût d'opportunité.
  - *Dans quelle mesure est-il possible de mettre en place un mécanisme optimal d'attribution des places en crèche ? Quelles sont les conséquences des différents choix de conception ?*
- **Puzzle n°2** : Parmi les populations les plus vulnérables on trouve une part importante de familles monoparentales bénéficiaires de minima-sociaux. Une solution proposée dans la droite ligne des politiques d'activation est l'accompagnement vers l'emploi.

[21] Voir Fiche N°1 de la COG.

– *Les programmes d'accompagnement global vers l'emploi permettent-ils aux familles monoparentales de sortir de la pauvreté par l'activité ?*

- **Puzzle n°3** : le système socio-fiscal français – particulièrement les aides des CAF – est complexe, peu lisible et les incitations très différentes suivant les configurations familiales
  – *Comment réagissent les mamans-solo pauvres aux incitations du système socio-fiscal français ?*

IV Cadre d'analyse

IV.1 Une approche evaluative centrée sur l'inférence causale

Comme le souligne simplement Joshua D. Angrist (2022) dans sa conférence à la remise de son prix Nobel - publiée ensuite dans Econometrica - "*Une stratégie empirique pour l'évaluation d'un programme ou d'une politique est un plan de recherche qui englobe la collecte de données, l'identification et l'estimation*". En reliant l'économétrie au monde des résultats potentiels hétérogènes introduits par Rubin (1974), les lauréats du Nobel 2021 ont ouvert la voie à la "*révolution de la crédibilité*", transformant rapidement la manière dont la plupart des économistes mènent leurs recherches, les questions qu'ils abordent et, surtout, la manière dont ils les abordent (Joshua D. Angrist and Pischke 2010).

Une constante dans cette thèse est l'attention portée à l'interprétation des analyses de données réalisées. J'introduis ici le modèle générique sur lequel se fondent la majorité de mes recherches.

Le modèle causal de Rubin Pour évaluer les effets d'une politique publique, comme celles que nous analysons dans cette thèse, il est nécessaire d'utiliser un groupe de comparaison pour représenter ce qui se serait produit pour les individus *traités*, et le principal défi est de déterminer si ce groupe de comparaison représente fidèlement le *contrefactuel*. En termes simples, il est difficile d'obtenir une interprétation causale des comparaisons statistiques si les individus traités et non traités ne sont pas "comparables" et/ou si les individus traités bénéficient plus ou moins que ce que les non traités auraient bénéficié.

La façon générale de représenter le problème s'appuie donc sur une expérience de pensée autour de *résultats potentiels*. L'intuition est la suivante:

Avant l'intervention, chaque famille évolue dans un contexte spécifique, caractérisé par son vécu, son histoire, ses préférences, ses compétences, etc. Au sein de la population, ces circonstances et attributs sont suffisamment riches pour être considérés comme des chocs individuels idiosyncratiques, *i.e*, aléatoires et indépendents les uns des autres. L'intervention agit elle aussi comme un "choc" dans cet équilibre et crée des potentiels de résultats, selon que la personne est traitée ou non. Une hypothèse fondamentale pour Rubin (1974) est que ces potentiels sont des attributs des personnes, des caractéristiques fixes et révélées par l'expérience. Formellement, pour chaque foyer i observé à une date t, il existe deux valeurs potentielles pour une variable de résultat Y : $Y_{it}(1)$ si elle est traitée et $Y_{it}(0)$ sinon. L'effet du programme au niveau individuel est la différence de résultats potentiels :

$$\delta_{it} = Y_{it}(1) - Y_{it}(0) \Leftrightarrow Y_{it}(1) = Y_{it}(0) + \delta_{it}$$

En postulant cette existence, on suppose concrètement que les résultats d'une personne ne sont pas influencés par ceux des autres (hypothèse de *stable unit treatment value assumption* - SUTVA) et que le "traitement" est homogène. Cependant, on n'observe jamais qu'une seule de ces valeurs potentielles, et le problème fondamental de l'inférence causale est que le contrefactuel est toujours inobservable. Il ne s'agit pas d'un problème de taille d'échantillon mais de données manquantes "dans l'univers". Toutefois, le statut de traitement (D_i) révèle le résultat potentiel :

$$Y_{it} = (1 - D_{it})Y_{it}(0) + D_{it}Y_{it}(1) \Leftrightarrow Y_{it} = Y_{it}(0) + D_{it}\delta_{it}$$

Avant que la politique ne soit mise en place, les résultats observés correspondent à $Y_{it} = Y_{it}(0)$ pour toutes les personnes de la population, et peuvent être analysés comme une variable aléatoire que l'on peut aussi écrire :

$$Y_{it} = Y_{it}(0) = E\left(Y_{it}(0)\right) + \varepsilon_{it}$$

Chaque observation peut être décomposée comme la somme de la moyenne de $Y_{it}(0)$ dans toute la population et d'un terme d'erreur idiosyncratique ε_{it} d'espérance nulle. C'est ce terme d'erreur qui résume en une dimension toute la complexité des individus, tandis que $E\left(Y_{it}(0)\right)$ résume la tendance dans la population. Si l'on considère des caractéristiques observables résumées dans une matrice **X**, qui peut notamment contenir ou correspondre au statut de traitement, la *loi des espérances itérées* permet aussi d'écrire:

$$Y_{it} = \mathbb{E}\Big[\mathbb{E}[Y_i|\mathbf{X} = \mathbf{x}]\Big] + \varepsilon_{it}$$

Lorsque la politique est mise en place, on peut décomposer la différence des espérances conditionnelles en fonction des groupes:

$$\begin{aligned}
\Delta &= \mathbb{E}[Y_{it}|D=1] - \mathbb{E}[Y_{it}|D=0] \\
&= \mathbb{E}[Y_{it}(1)|D=1] - \mathbb{E}[Y_{it}(0)|D=0] \\
&= \underbrace{\mathbb{E}[Y_{it}(1) - Y_{it}(0)|D=1]}_{ATT} - \underbrace{(\mathbb{E}[Y_{it}(0)|D=1] - \mathbb{E}[Y_{it}(0)|D=0])}_{\text{Bias de sélection}}
\end{aligned}$$

La première ligne souligne, sous *SUTVA*, que les données observées des différents groupes révèlent leurs résultats potentiels. La seconde ligne s'obtient en ajoutant et soustrayant le contrefactuel du groupe traité, ce qui permet de décomposer l'écart moyen entre traités et non-traités dans la population en deux paramètres interprétables :

- L'effet moyen du programme sur les traités mesure la moyenne des effets individuels du traitement parmi la population traitée ;
- Le biais de sélection correspond aux différences moyennes entre traités et non-traités si la politique n'existait pas.

On a donc une expression pour le biais de sélection qui s'interprète facilement. Dans la situation où le programme n'existerait pas, celles qui auraient participé auraient en moyenne des résultats différents de celles qui n'auraient pas participé.

Avec une manipulation similaire, on montre également facilement:

$$\Delta = \underbrace{\mathbb{E}[Y_{it}(1) - Y_{it}(0)]}_{ATE} + \underbrace{\mathbb{E}\Big[Y_i(0)|D_{it}=1\Big] - \mathbb{E}\Big[Y_{it}(0)|D_i=0\Big]}_{\text{Bias de sélection}} + \underbrace{(1-\pi)(ATT - ATU)}_{\text{Hétérogénéité de l'effet}}$$

Où π correspond à la proportion d'individus traités.

Plus intuitivement, on s'attend à ce que les individus qui *choisissent* un programme ou bénéficient d'une politique puissent être différents de ceux qui *choisissent de ne pas* y participer. Ils peuvent être plus disponibles, plus motivés, ou avoir d'autres caractéristiques qui les distinguent. Ce phénomène de sélection fait que dans l'ensemble, les groupes traités et non-traités ne sont pas composés d'individus semblables, et donc les comparer intègre ces différences de composition, affectant ainsi les résultats moyens.

Dit autrement, en présence de biais de sélection, les résultats des non-traités ne représentent pas bien ce qui se serait passé pour les traités.

La formule du biais de sélection nous donne aussi une condition pour une interprétation causale :

$$E(Y_{it}(0)|D_{it}=1) = E(Y_{it}(0)|D_{it}=0)$$

Cette condition est en particulier vérifiée si la participation est indépendante des résultats potentiels, ce qui est notamment le cas lorsque la participation est aléatoirement assignée et que tout le monde se conforme à l'assignation.

Expérimentations aléatoires de terrain La randomisation est le processus qui consiste à rendre quelque chose aléatoire. Un processus aléatoire est une séquence de variables aléatoires décrivant un processus dont les résultats ne suivent pas un modèle déterministe, mais qui peut être décrit par des distributions de probabilité. Par exemple, un échantillon aléatoire d'individus d'une population se réfère à un échantillon où chaque individu a une probabilité connue d'être échantillonné. Cette probabilité d'échantillonnage sert à déduire des valeurs plausibles des paramètres de l'ensemble de la population à l'aide des statistiques de l'échantillon.

Un échantillon aléatoire de grande taille garantit la représentativité de la population et des estimations plus précises, c'est-à-dire une variabilité d'échantillonnage moindre. La variabilité d'échantillonnage d'une estimation est une mesure de la variation de l'estimation d'un échantillon à l'autre.

Les expériences aléatoires sont des *loteries* qui répartissent de manière aléatoire les sujets d'un échantillon dans des groupes de recherche, chacun d'entre eux se voyant offrir un traitement différent. L'assignation aléatoire assure l'équilibre entre les groupes, ce qui garantit l'absence de différences *ex-ante*, en espérance.

L'échantillon d'une expérience aléatoire peut, ou non, être un échantillon aléatoire. Il s'agit de deux questions distinctes. La randomisation garantit la *validité interne*, c'est-à-dire l'absence de biais dans l'estimation des effets du traitement dans cet échantillon. Les expériences aléatoires de grande envergure donnent des estimateurs plus précis, mais la déduction pour une population plus large dépend du processus d'échantillonnage, de la sensibilisation des participants, du moment et du lieu spécifiques, etc. Cela renvoie à la question de la *validité externe*, c'est-à-dire à la généralité de ces résultats.

La méthode expérimentale généralement attribuée à Fischer (1935) et Neyman (1934) et résout immédiatement ce problème en assurant l'indépendance (conditionnelle) des résultats potentiels. Supposons que le décideur politique randomise l'accès au traitement et que $Z_i = \mathbb{1}(\text{Assigned treatment})$. Dans la plupart des cas, l'affectation aléatoire est conditionnée par certains attributs **X** et donne $Y_i(1), Y_i(0) \perp Z_i|\mathbf{X}$. En cas de conformité parfaite, $Z_i = D_i$ et

$$\mathbb{E}[Y_i(0)|D=1] - \mathbb{E}[Y_i(0)|D=0] = 0$$

Dans ce cas, la simple différence de moyenne entre ceux qui ont reçu le traitement et ceux qui ne l'ont pas reçu permet d'estimer l'effet moyen du traitement sur les traités, ou l'effet moyen du traitement si l'échantillon est représentatif de la population.

En cas de conformité imparfaite, nous devons changer les notations et ajouter une hypothèse supplémentaire pour utiliser des variables instrumentales. Maintenant, le statut du traitement est également considéré comme un résultat potentiel et nous écrivons:

$$\begin{aligned} D_i = & \begin{cases} D_{1i} = \text{i's treatment status when } Z_i = 1 \\ D_{0i} = \text{i's treatment status when } Z_i = 0 \end{cases} \\ = & D_{0i} + (D_{1i} - D_{0i})Z_i \\ = & \pi_0 + \pi_{1i} Z_i + \zeta_i \end{aligned}$$

La notation précédente nous donne une équation de participation (*first stage*) permettant à l'instrument Z d'avoir des effets différents entre les individus. Une première condition est que l'effet moyen soit suffisamment fort pour que $\mathbb{E}[D_{1i} - D_{0i}] \neq 0$.

Avec ces notations, nous pouvons diviser la population en quatre groupes en fonction de leur réaction à l'instrument:

- Les "toujours preneurs" (*always-takers*) ont $D = 1$ peu importe la valeur de Z
- Les "jamais preneurs" (*never-takers*) ont $D = 0$ peu importe la valeur de Z
- Les "mobilisables" (*compliers*) ont une valeur de D qui suit la valeur potentielle de Z
- Les "défiants" (*defiers*) ont une valeur de D opposée à la valeur réalisée de Z

Pour identifier l'effet de la politique permettant des effets de traitement hétérogènes, il ne doit pas y avoir de défiants, c'est-à-dire d'individus qui font le contraire de ce qui leur a été assigné. On suppose que si l'instrument n'a pas d'effet sur certains individus, tous ceux qui sont affectés le sont dans la même direction. Formellement:

$$\pi_{1i} \geq 0 \quad | \quad \pi_{1i} \leq 0 \quad \forall i$$

Enfin, nous devons exclure que l'affectation Z ait d'autres effets sur le résultat que par l'intermédiaire de la variable de traitement D. Cette hypothèse est appelée *restriction d'exclusion* et restreint la combinaison des résultats potentiels $Y_{it}(d,z)$ comme étant uniquement une fonction de d:

$$Y_{it}(d,1) = Y_{it}(d,0) \equiv Y_{di} \text{ pour } d = 0,1.$$

Nous pouvons donc réécrire les résultats observés comme étant une fonction de ces résultats potentiels :

$$\begin{aligned} Y_{it} &= Y_(0,Z_i) + (Y_{it}(1,Z_i) - Y_{it}(0,Z_i))D_i \\ &= Y_{0i} + (Y_{1i} - Y_{0i})D_i \\ &= \alpha + \rho_i D_i + \varepsilon_i \end{aligned}$$

Avec $\mathbb{E}[Y_{0i}] = \alpha$ et $(Y_{1i} - Y_{0i}) = \rho_i$ l'effet individuel du traitement.

En résumé, si les

- **Independence** $\left(Y_{it}(D_1,1), Y_{it}(D_0,0), D_{1i}, D_{0i}\right) \perp Z_i$
- **exclusion** $Y_{it}(d,1) = Y_{it}(d,0) \equiv Y_{di}$ pour $d = 0,1.$
- **first-stage** $\mathbb{E}[D_{1i} - D_{0i}] \neq 0$
- **monotonicité** $D_{1i} - D_{0i} \geq 0 \mid D_{1i} - D_{0i} \leq 0 \; \forall \; i$

Dans ces conditions, le théorème principal de J. D. Angrist and Imbens (1995) est que le ratio WALD estime le **local average treatment effect**(LATE) :

$$\begin{aligned} \frac{\mathbb{E}[Y_{it}|Z_i = 1] - \mathbb{E}[Y_{it}|Z_i = 0]}{\mathbb{E}[D_i|Z_i = 1] - \mathbb{E}[D_i|Z_i = 0]} &= \mathbb{E}[Y_{1i} - Y_{0i}|D_{1i} > D_{0i}] \\ &= \mathbb{E}[\rho_i|\pi_{1i} > 0] \end{aligned}$$

Le ratio de Wald ou les coefficients 2SLS sont des estimations cohérentes du LATE : l'*effet causal moyen sur les mobilisables*, et eux seuls. Les mobilisables sont le seul groupe dont les unités ont été observées dans les deux traitements (étant donné que les défiants ont été exclus). Le théorème LATE et les nombreux nouveaux résultats qui tournent autour tissent la trame de fond de cette thèse.

De façon générale, l'évaluation causales des politiques implique une stratégie empirique comprenant trois éléments clés:

1) **Identifier** un lien de cause à effet: sous quelles hypothèses peut-on identifier des paramètres causaux ?
2) **Estimer**: À partir d'un échantillon de données, comment estimer les paramètres d'intérêt identifiés dans l'étape précédent ?
3) **Inférer** au dela de cet échantillon: Caractériser la précision des estimations et discuter la validité au delà de ce jeu de données.

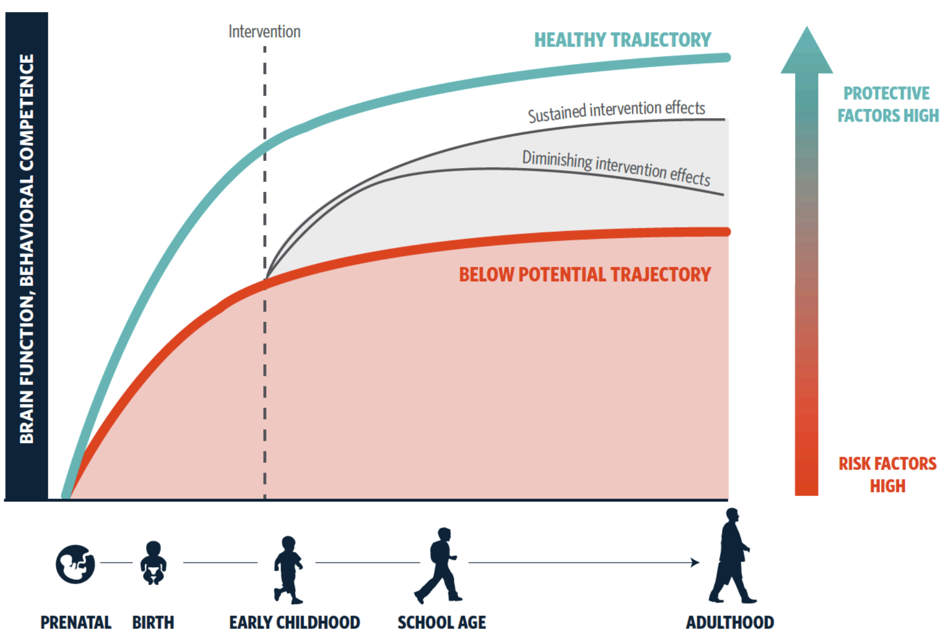
À chaque challenge sa source de bias:

- L'identification pose la question du biais de sélection et de l'interprétation causale, mais ce n'est pas le seul ;
- Différents estimateurs peuvent produire des estimations plus ou moins éloignées des paramètres de la population, suivant notamment la taille de l'échantillon ou de propriétés de convergence asymptotique.
- En fonction de comment les données sont produites, du nombre d'observation, du nombre de paramètres testés, nous pouvons plus ou moins facilement interpréter les résultats pour une population plus générale et tester les estimateurs contre des valeurs "*a priori*", et notamment contre un effet nul.

IV.2 Vers le design et l’évaluation de politiques sociales et familiales

Le potentiel des interventions précoces Il est largement reconnu dans le monde académique que la prise en charge de la petite enfance joue un rôle crucial dans les trajectoires de vie des enfants et des parents, avec des implications profondes pour le bien-être de la société, l’égalité des genres et les résultats économiques. De manière simplifiée, les jeunes enfants développent un système neuronal très malléable, les rendant vulnérables aux influences environnementales et capables de bénéficier d’interventions (Case, Fertig, and Paxson 2005; Howard-Jones, Washbrook, and Meadows 2012; Jackson, Kiernan, and McLanahan 2017). Le développement de l’enfant est déterminé par l’intégrité et le fonctionnement du système nerveux central, ainsi que par des facteurs environnementaux positifs et négatifs qui affectent ce développement. Ce processus débute très tôt dans la vie, dès la gestation, et se poursuit tout au long de la petite enfance, avec des rythmes de développement différents selon les zones du cerveau. Le développement rapide du cerveau et du corps, combiné aux interactions sociales et à l’environnement, contribue à des évolutions particulièrement rapides des capacités des enfants au cours de cette période (Fernald et al. 2017). Les facteurs environnementaux positifs englobent divers éléments tels que les conditions de vie et de santé, la nutrition, la sécurité, les soins, les interactions sociales et l’apprentissage précoce. Les interventions visant un ou plusieurs de ces facteurs environnementaux peuvent influencer le développement des enfants et les orienter vers des trajectoires de vie plus favorables, voire durables [voir Figure 5].

Figure 5: Les interventions précoces pour améliorer les trajectoires de vie des enfants d’après Fernald et al. (2017)



Les premières étapes de la vie, de la petite enfance à l’enfance, sont des périodes cruciales qui déterminent largement les trajectoires futures des individus (E. D. Bono et al. 2016; Francesconi and Heckman 2016; Anderson et al. 2021; Grey et al. 2022). L’investissement parental et les ressources économiques disponibles jouent un rôle déterminant dans ces trajectoires. Les politiques publiques axées sur la petite enfance offrent ainsi une opportunité stratégique en permettant d’agir sur plusieurs fronts simultanément.

Tout d’abord, le temps passé par les enfants dans des environnements de qualité, tels que les structures d’accueil de la petite enfance, exerce une influence significative sur leur développement cognitif et socio-émotionnel (Case, Fertig, and Paxson 2005; Flavio Cunha et al. 2006; Camilli et al. 2010; Burger 2010; Nores and Barnett 2010; R. C. Johnson and Schoeni 2011; Barnett 2011; Currie and Almond 2011; O. P. Attanasio 2015; Walters 2015; Elango et al. 2016; Kholoptseva 2016; Carbuccia et al. 2020). Ensuite, ces politiques contribuent à alléger les contraintes logistiques et financières auxquelles sont confrontés les parents, en particulier les femmes, favorisant ainsi leur intégration ou leur maintien sur le marché du travail (Kimmel 1998; Nollenberger and Rodríguez-Planas 2015; Morrissey 2017; Simintzi, Xu, and Xu 2022).

Cependant, malgré l'enthousiasme autour de ces politiques, leur évaluation ne donne pas toujours les résultats escomptés (Cantillon 2011; Ghysels and Van Lancker 2011; Vandenbroucke and Vleminckx 2011a; G. J. Duncan et al. 2012a). Pour comprendre ces divergences, les chercheurs se sont penchés sur les éléments constitutifs des politiques d'accueil de la petite enfance, notamment la qualité des interactions, l'environnement offert aux enfants, ainsi que les caractéristiques des professionnels encadrants.

L'importance de la qualité des interventions Tout d'abord, la qualité des services de garde d'enfants revêt une importance cruciale, à la fois en termes absolus et par rapport à d'autres options de garde (Marshall 2004; Manning et al. 2017; Blanden et al. 2022). Lorsque les services d'accueil offrent un environnement stimulant durant les premières années de vie, cela peut favoriser un développement holistique, dont les effets peuvent être durables. Cependant, en l'absence d'une qualité suffisante dans l'accueil, les soins, les interactions, etc., certains enfants peuvent ne pas bénéficier de ces effets positifs, voire subir des conséquences négatives (N. D. Gupta and Simonsen 2010; Felfe, Nollenberger, and Rodríguez-Planas 2015; Havnes and Mogstad 2015; Gansen 2017; Felfe and Lalive 2018; Ichino, Fort, and Zanella 2019; Kulic et al. 2019; Rabe 2019).

De plus, le développement du capital humain est un processus auto-renforçant avec des complémentarités dynamiques (F. Cunha and Heckman 2007). En d'autres termes, les compétences les plus complexes émergent des compétences de base déjà acquises, et les compétences acquises avant un nouvel investissement dans le capital humain augmentent la productivité de cet investissement (Lubotsky and Kaestner 2016; R. Johnson and Jackson 2017). Ainsi, maintenir et accroître les bénéfices des interventions précoces nécessite un investissement éducatif continu tout au long de l'enfance et au-delà.

Cependant, investir dans le capital humain ne se résume pas à une simple augmentation des dépenses éducatives. L'investissement doit entraîner un changement, et ce changement doit présenter des caractéristiques spécifiques. En France, une étude menée par C. de Chaisemartin et al. (2021) a réalisé une expérimentation aléatoire contrôlée à grande échelle pour évaluer un programme de formation destiné aux professionnels de la petite enfance, visant à améliorer la quantité et la qualité des interactions langagières avec les bébés. Toutefois, cette étude a eu du mal à observer des changements significatifs dans les résultats des enfants. Les chercheurs discutent des explications plausibles pour ces résultats décevants, suggérant notamment que le programme a eu peu d'effets initiaux sur les pratiques des professionnels.

Selon Bailey et al. (2017), des effets persistants des interventions précoces nécessitent une action sur des compétences, des comportements et des capacités qui partagent trois caractéristiques. Ces interventions doivent être :

- suffisamment malléables pour être affectés par l'intervention,
- essentiels pour le succès, et surtout,
- en l'absence d'intervention, ils ne doivent pas se développer spontanément (beaucoup moins, ou différemment).

van Huizen and Plantenga (2018) ont entrepris une synthèse systématique de la littérature quasi-expérimentale portant sur les effets de l'accès aux modes d'accueil entre 2005 et 2017, couvrant l'ensemble des pays développés. Leur analyse, basée sur 250 estimations issues de 30 études, a exploré une gamme variée de résultats relatifs au développement des enfants, allant du développement cognitif et non cognitif pendant la petite enfance aux résultats éducatifs et aux revenus à l'âge adulte. Les conclusions de cette méta-analyse révèlent une série de résultats nuancés. L'âge d'entrée dans les modes d'accueil ne semble pas être un facteur déterminant, tandis que l'intensité de l'accueil peut jouer un rôle dans certains cas. Les services publics de la petite enfance semblent avoir des effets plus significatifs sur le développement des enfants que les services privés. Cependant, le résultat le plus saillant est que les effets positifs sont principalement concentrés parmi les enfants issus de familles d'origine sociale défavorisée.

Inégalités d'accès aux modes d'accueil: *l'effet Matthieu* Pourtant, la sous-population des familles utilisant les services de garde d'enfants est fortement sélectionnée en faveur des familles à revenus élevés (Lancker and Ghysels 2012; Petitclerc et al. 2017). Ce phénomène, connu dans la littérature sur l'investissement social sous le nom d'effet "Matthieu"[22], soulève des questions importantes avec des implications politiques diverses.

L'effet Matthieu peut-il être attribué à des facteurs liés à la demande ? Est-il influencé par la démographie locale, les préférences des parents, ou les normes culturelles relatives à la parentalité ? Ou est-ce plutôt dû à des contraintes du côté de l'offre, telles que des capacités limitées, des coûts élevés, ou un cadre institutionnel contraignant et dissuasif ? Jusqu'à présent, les données suggèrent que les contraintes liées à l'offre jouent un rôle plus prépondérant dans l'explication de l'effet Matthieu que les facteurs liés à la demande (Farfan-Portet, Lorant, and Petrella 2011; Abrassart and Bonoli 2015 ; Pavolini and Van Lancker 2018 ; Carbuccia, Thouzeau, et al. 2023).

L'une des conséquences de cette sélection endogène est la variation considérable de l'impact de l'accès à la garde d'enfants sur le marché du travail, les inégalités de genre, ou le développement de l'enfant, selon le contexte, les instruments de politique publique utilisés, le milieu socio-économique des parents, et la structure familiale (Misra, Moller, and Budig 2007; Tekin 2007; Farfan-Portet, Lorant, and Petrella 2011; Nollenberger and Rodríguez-Planas 2015; Kottelenberg and Lehrer 2016; Zoch 2020; Bousselin 2022). Cette variation dépend également de l'existence d'autres modes d'accueil alternatifs, d'institutions du marché du travail telles que des horaires de travail flexibles, des compléments aux bas salaires, et d'autres politiques actives du marché du travail (Gorey 2009; Fagnani 2010; Fleche, Lepinteur, and Powdthavee 2018b).

Les politiques d'accueil du jeune enfant interagissent avec d'autres politiques sociales telles que les prestations familiales, le congé parental, les politiques actives du marché du travail, les politiques fiscales et redistributives, les politiques d'éducation ou de santé, etc. (Olivetti and Petrongolo 2017; D. Gupta, Jessen, and Jonas 2023). La générosité des politiques d'accueil du jeune enfant est positivement corrélée à la part de femmes élues au parlement (Giuliano Bonoli and Reber 2010). De plus, les normes sociales et les rôles genrés assignés aux hommes et aux femmes affectent également l'efficacité des politiques petites enfances. Ces normes peuvent être internalisées par les parents mais aussi par les professionnels de la petite enfance et les décideurs (Holloway 1998; Windebank 2001; N. M. Fortin 2005; Blau and Kahn 2017; Juhn and McCue 2017; Chung 2020; Cavapozzi, Francesconi, and Nicoletti 2021; H. Kleven et al. 2023; Briselli and Gonzalez 2023).

Au sein des diverses formes que prennent les États providences modernes, qui diffusent et renforcent les différents points de vue sur le rôle des mères et des pères dans la société, certaines combinaisons de politiques peuvent limiter l'impact positif de la garde d'enfants sur la participation au marché du travail et la réduction des inégalités (Van Lancker and Ghysels 2016; Ünver, Bircan, and Nicaise 2018; Pavolini and Van Lancker 2018).

Par conséquent, pour obtenir des effets positifs importants sur la société, l'élaboration de bonnes politiques d'accueil du jeune enfant nécessite une combinaison adéquate de politiques actives en matière de marché du travail, de garde d'enfants et de soutien aux familles (A. Hemerijck and Huguenot-Noël 2022).

On trouve ainsi dans la littérature sensible aux enjeux de causalité des éléments de réponses importants pour la définition de politiques sociales et familiales. L'investissement social offre ainsi un cadre normatif pour évaluer les politiques menées.

IV.3 Les projets menées pour réaliser cette thèse

Investissement social dans l'accueil du jeune enfant: le projet ISAJE Le projet ISAJE (Investissement Social dans l'Accueil du Jeune Enfant) prend la forme d'une collaboration scientifique avec Julien Combe et d'un partenariat institutionnl entre la Caisse Nationale des Allocations Familiales (Cnaf) et l'École Polytechnique (X). La finalité de ce projet est de pallier le manque de données probantes concernant les effets de l'accès à des modes d'accueil formels sur les familles.

Lorsque l'on examine les répercussions d'une place en crèche, il est crucial de se demander ce qui aurait pu se produire pour une même famille si elle n'avait pas bénéficié de cette opportunité. Dans la plupart des études existantes, l'accès aux données se fait bien après les attributions, ce qui conduit à utiliser des familles sans place en crèche mais

[22] Ou *Matthew effect*, selon l'expression attribué à Robert K. Merton en référence à l'extrait de l'évangile selon Saint Matthieu « *Car on donnera à celui qui a, et il sera dans l'abondance, mais à celui qui n'a pas on ôtera même ce qu'il a.* »"

présentant des caractéristiques statistiquement similaires à celles ayant obtenu une place, afin de représenter ce qui aurait pu se produire autrement (Gomajee et al. 2017, 2018; Grobon, Panico, and Solaz 2019). Cependant, cette hypothèse d'équivalence entre les deux groupes est souvent difficile à vérifier.

La compréhension des mécanismes qui mènent certaines familles à obtenir une place en crèche et d'autres non remet en question la validité de ces comparaisons. En effet, la simple démarche de candidature pour une place en crèche distingue déjà les familles, et les facteurs influençant ces décisions sont souvent complexes à mesurer. Ils peuvent être liés à l'attitude des parents, à leurs interactions avec leur enfant, aux contraintes familiales ou contextuelles, entre autres. De plus, parmi les parents sollicitant une place en crèche, ceux qui réussissent à l'obtenir diffèrent généralement de ceux qui n'y parviennent pas, en raison notamment des critères établis par les autorités locales. Ces critères variant d'une commune à l'autre, les chercheurs ont souvent peu d'accès à ces informations.

Quelques rares travaux ont recours à des expériences naturelles, exploitant les variations d'accès aux modes d'accueil entre différentes familles, indépendamment des résultats potentiels autres que leur effet sur l'accès en crèche. Par exemple, Berger, Panico, and Solaz (2021) ont utilisé les différences d'accès entre les enfants nés au printemps et ceux nés en automne, ainsi que les disparités entre les villes présentant des taux de couverture des modes d'accueil formel plus ou moins élevés. Leurs résultats ont montré que l'accès à une place en crèche favorise le développement langagier à 2 ans, sans impact sur la motricité mais avec des effets négatifs sur les mesures de comportement. De plus, ces effets positifs sur le langage sont principalement concentrés parmi les enfants issus de milieux défavorisés. Par ailleurs, Pora (2020) a comparé l'emploi des mères dans des zones où l'offre de crèche a fortement augmenté à celui dans des zones où l'offre est restée constante, sans observer d'effet moyen sur l'emploi.

Cependant, malgré ces avancées, le domaine de la recherche reste encore insuffisamment exploré, d'autant plus que le système d'accueil du jeune enfant en France se distingue fortement des autres pays de l'OCDE.

Le projet de recherche a été lancé au début de l'année 2019, et la première année a été consacrée à la définition de l'intervention et à la prospection des terrains d'expérimentation. Grâce à un groupe de travail composé d'élus, de directeurs de services petite enfance, d'agents de la Caisse d'Allocations Familiales (Caf) et de chercheurs, les contraintes de faisabilité du projet ont pu être définies. À la fin de l'année, plusieurs territoires ont exprimé leur intérêt pour cette démarche, mais la tenue des élections municipales l'année suivante a empêché la signature de contrats avec les communes. De plus, la crise sanitaire est venue perturber davantage le déploiement de l'expérimentation, laissant une empreinte durable sur l'ensemble du secteur de la petite enfance.

Nous avons pu lancer une première phase pilote en 2020 à Valence-Romans Agglo. En tant que communauté d'agglomération, le report des élections n'a pas eu d'impact sur notre capacité à contractualiser, et l'automatisation des commissions d'attribution s'est avérée particulièrement utile pendant le confinement. Grâce à ce premier essai réussi, de nouveaux territoires ont progressivement accepté de rejoindre le projet.

Cependant, pour pouvoir mesurer les effets avec une précision suffisante, nous avions besoin d'un nombre important de parents, et nous avons rencontré des difficultés à recruter suffisamment de territoires pour constituer un échantillon de taille adéquate pour atteindre le niveau de précision statistique souhaité. Par conséquent, l'objectif initial de tester en conditions cliniques tous les enfants a dû être abandonné. Néanmoins, l'ambition du projet ISAJE n'a pas été réduite, mais au contraire, elle a été renforcée.

Dans un premier temps, l'automatisation des commissions d'attribution et l'implication de Julien Combe nous ont permis de proposer aux territoires des procédures flexibles et adaptées à chaque contexte, dotées de propriétés attrayantes. La phase I du projet a consisté à programmer puis utiliser ces outils pour proposer des affectations dans les territoires participants.

Ensuite, la phase II, que nous amorçons maintenant, repose sur l'appariement des données des commissions d'affectation avec les bases de données des allocataires de la Cnaf, auxquelles nous ajoutons les données des parents d'enfants de moins de 3 ans résidant dans ces territoires. Ces données sont également géolocalisées et contiennent un ensemble riche d'informations sur les ménages, leurs caractéristiques, leurs ressources, leurs aides, leurs configurations, etc.

Enfin, l'objectif initial reste inchangé. L'appariement avec les données de la Cnaf nous fournit des informations précieuses sur les revenus des familles, ce qui nous permettra de mesurer les effets de l'obtention d'une place en crèche sur l'activité des parents, leurs ressources, l'évolution de leur configuration familiale, et de mener des analyses différenciées selon différentes caractéristiques. De plus, la mise en place de ce suivi n'exclut pas, à terme, d'organiser des enquêtes pour mesurer les compétences des enfants à un stade ultérieur, soit directement soit par appariement avec les données du ministère de l'Éducation nationale.

Investissement social et activation: l'évaluation du programme Reliance Les femmes chefs de famille monoparentale bénéficiaires du RSA depuis longtemps vivent souvent dans des conditions précaires. Dans le cadre d'un programme initié sur le territoire du Grand Nancy en Meurthe-Et-Moselle, nous avons été sollicités pour proposer un design de recherche permettant d'évaluer cette initiative d'accompagnement global intensif ciblant spécifiquement cette population.

Depuis 2017, le conseil départemental de Meurthe-et-Moselle, la Caisse d'Allocations Familiales (Caf) locale et la Caisse nationale ont développé un programme visant à soutenir durablement les femmes chefs de famille monoparentale bénéficiaires du RSA dans leur réinsertion sociale et professionnelle. Ce projet concerne un public résidant dans la métropole du Grand Nancy, une zone urbaine densément peuplée bénéficiant d'un réseau de transports en commun développé et d'une gamme variée de services de proximité, notamment en matière de garde d'enfants (structures petite enfance, périscolaires). Baptisé "Reliance", ce dispositif expérimental est mis en œuvre par trois associations locales bien établies et expérimentées : Arélia, Ulis et Ecoval. L'accompagnement d'une durée d'un an comprend à la fois des séances collectives et individuelles et s'appuie sur un cadre conceptuel inspiré de recherches pluridisciplinaires et de l'expertise des porteurs de projet. Les travailleurs sociaux sont qualifiés et expérimentés, et les porteurs de projet bénéficient d'une bonne connaissance du territoire.

Les hypothèses concernant les effets du programme reposent sur une approche capacitaire, où les participantes acquièrent ou renforcent des compétences et des réseaux, ainsi que sur une approche émancipatrice, visant à lever certains obstacles spécifiques à chaque famille. L'objectif de Reliance est donc d'assurer une réinsertion durable des bénéficiaires dans le monde du travail, dans une perspective d'investissement social, en renforçant leurs capacités à s'intégrer pleinement dans la société et le marché du travail. On espère ainsi, par son efficacité, éviter des coûts supplémentaires pour les finances publiques. Les porteurs de projet ont mandaté la Cnaf pour réaliser une évaluation d'impact économétrique en parallèle d'une évaluation qualitative financée par le Conseil départemental et réalisée par le bureau d'études FORS Recherche sociale (FORS 2020).

Le programme Reliance et cette évaluation présente plusieurs spécificités qui nous permettent de le relier à d'autres expériences analysées dans la littérature:

1) **Un public spécifique**: les familles monoparentales bénéficiaires de minima-sociaux depuis au moins 2 ans
2) Un accompagnement social **délégué à des opérateurs privés** (associatifs)
3) Un accompagnement **intensif long à la fois collectif et individualisé**
4) Un programme comparables aux dispositifs *welfare-to-work* de part son inscription dans le champ des "droits et devoirs" qui lui confère un caractère **perçu comme obligatoire** et où la peur de sanctions permet une assez forte mobilisation.
5) Une évaluation qualitative et quantitative basée sur une expérimentation aléatoire contrôlée.

Le projet était conçu pour accompagner une centaine de personnes par an pour initialement 3 cohortes. Depuis, il a été décidé de poursuivre avec deux nouvelles cohortes en 2021 et 2022, en raison de la crise Covid-19 qui a perturbé l'accompagnement de la cohorte 2020 et possiblement les perspectives d'insertion de la cohorte 2019 qui a quitté l'accompagnement juste au début du premier confinement. De plus, dans l'attente des résultats d'évaluation il apparaissait préférable de continuer le programme et de décider ou non de sa poursuite.

Des rapports officiels récents ont été très critiques sur la mise en œuvre du suivi et de la mise en oeuvre de l'accompagnement social obligatoire pour les allocataires du RSA (Pitollat and Klein 2018, Aout ; Damon 2018 ; Cour des comptes 2022). Cette expérience emprunte une voie totalement différente avec un soutien intensif, un processus de recrutement innovant et une forte insistance sur la mesure de ses effets. Elle a également bénéficié d'un soutien politique et médiatique important, avec notamment une visite officielle[23] de la ministre des Solidarités et de la Santé, Agnès Buzyn, et de la secrétaire d'État associée, Christelle Dubos.

L'expérimentation de Reliance a été pré-enregistrée sur le site www.socialscienceregistry.org ; expérimentation n°5930 avant d'accéder aux données. Dans cet enregistement sont décrits en avance les choix méthodologiques qui

[23] Meurthe-et-moselle.fr Actu - visite ministre

guident le Chapitre 2. De plus, cette expérience est approuvée et suivie par l'*International review board* de l'École d'Économie de Paris (PSE) qui garantit que ces travaux respectent l'éthique de la recherche sur sujet humains.

V Contributions du Chapitre 1

Ce chapitre est co-signé avec Julien Combe, professeur d'économie à l'école polytechnique et porte sur les mécanismes d'affectations des places en crèche et une analyse des inégalités sociale d'accès.

V.1 Lier théorie, application et évaluation

Dans cet article, nous considérons l'accès à la crèche comme un problème d'appariement où la demande rencontre l'offre grâce à un mécanisme d'attribution centralisé organisé par les autorités locales. Sur la base de la définition du problème par les décideurs politiques, nos principales questions de recherche sont les suivantes :

Dans quelle mesure un mécanisme d'attribution optimal pour la crèche est-il possible ? Quelles sont les conséquences des différents choix de conception ?

Nos principales contributions sont doubles. Premièrement, nous fournissons des modèles et des algorithmes pour différentes versions du problème d'attribution de crèche avec des propriétés bien comprises. Les applications fournissent aux décideurs de nouveaux outils pour attribuer aux parents leur crèche préférée suivant des priorités avec diversité et des contraintes multidimensionnelles. Deuxièmement, nous utilisons deux études de cas pour analyser les effets des contraintes de diversité, des priorités et des mécanismes d'attribution sur les inégalités dans l'accès en crèche selon différentes dimensions.

Nous avons remarqué que l'organisation des commissions d'affectation présentait certaines caractéristiques familières. En particulier:

i) les affectations sont centralisées par les collectivités locales,
ii) les parents s'inscrivent tout au long de l'année et soumettent leurs préférences pour les crèches qu'elles préfèrent,
iii) la plupart des offres sont disponibles en septembre lorsque les enfants plus âgés passent à l'école maternelle, et
iv) les collectivités locales organisent un comité d'attribution principal au printemps pour attribuer ces places.

Le problème d'attribution des places en crèche est donc très similaire au problème de choix d'école pour lesquels des solutions bien définies et de nombreuses applications réussies existent déjà (Abdulkadiroğlu and Sönmez 2003).

Dans cette littérature, l'objectif est généralement de définir des appariements *stables* et des algorithmes pour les trouver. La *stabilité* est une propriété que nous voulons qu'un appariement respecte afin de justifier aux parents *pourquoi* l'attribution est ainsi. La motivation théorique pour se concentrer sur des ensembles d'appariements stables est que si le résultat du marché est instable, il y a des agents ou des paires d'agents qui ont l'incitation à contourner l'appariement (A. E. Roth 2002). Dans certains cas, les chercheurs et/ou les décideurs utilisent déjà les résultats de la littérature sur la conception des marchés (*Market design*) pour fournir des mécanismes d'attribution des places en crèche (voir notre revue à la Section II du Chapitre 1).

Le terme *marché* a une charge politique importante, mais pas pour tout le monde et pas de la même manière. Pour certaines personnes, le terme de *marché* n'évoque rien de particulier, c'est un terme *inerte* qui désigne, notamment, un lieu où l'on fait des achats. Pour d'autres, il évoque le vocabulaire des *néolibéraux* et sa simple prononciation suscite une forme de défiance envers celui ou celle qui l'emploie. Pour des chercheurs en économie, le terme *marché* a un sens bien particulier. De plus, les *marchés* que nous définissons équilibrent offre et demande à partir d'un *algorithme*. Là aussi, la polysémie du terme et les enjeux politiques associés sont grands. Celà nous invite donc à définir clairement ces notions ou, du moins, préciser le sens que nous leur conférons dans ces travaux.

La recette des affectations Dans notre définition, un algorithme est *une suite d'instructions logiques* pour atteindre un objectif défini à l'avance. Pour préciser cette sémantique, on peut faire une métaphore culinaire: un algorithme est une recette qui réussi à tous les coups. En cuisine, il y a un problème : faire un plat. Pour cela, il faut choisir un plat, ce qui détermine la recette à suivre. Le plat a des propriétés (goût, texture etc.) et chaque étape de la recette contribue à les obtenir. Cette dernière contient les ingrédients - éléments du problème - et une succession de tâches précises documentant la façon de les préparer, les manipuler, etc. En suivant précisément ces instructions avec un panier d'ingrédients, la recette permet d'atteindre l'objectif et le plat est réussi, il a les caractéristiques attendues, et tout le monde se régale. En reproduisant la recette avec un autre panier d'ingrédients, on obtient le même plat, avec les mêmes caractéristiques.

Pour nous, le problème est d'affecter les places en crèche, des choix politiques conduisent à définir l'objectif et la recette qui peut permettre de l'atteindre. Une recette contient les éléments du problème (Les demandes avec les vœux des familles, les priorités définies par le territoires, les crèches avec leurs capacités, parfois des places réservés pour des enfants de différents types etc.) et les étapes de traitement. Ces instructions sont claires et, puisqu'elles sont logiques, les tâches peuvent facilement être confiées à un ordinateur[24]. Néanmoins, ces instructions peuvent très bien être suivies par une personne ou une commission. Ce n'est rien d'autre qu'une fiche de procédure traduite dans un langage interprétable par un calculateur. Le principal intérêt de l'automatisation est le gain de temps qu'il occasionne et l'absence d'erreur ou de déviation. C'est pourquoi, nous parlons souvent de *procédure automatique*, expression qui caractérise mieux ce que sont nos algorithmes. En particulier, ils ne fonctionnent pas du tout comme une "intelligence artificielle". Il n'y a pas d'apprentissage statistique (*machine learning*).

Un algorithme est donc une *procédure* ϕ qui exécute la même succession de tâches à partir d'un *problème* P posé, et composé d'une série d'éléments qui dépendent du problème. Cette procédure retourne une affectation μ. Si les éléments d'un problème P et P' sont les mêmes, l'affectation retournée est la même. Elle est déterministe[25].

Pour résumé, nous définissons et utilisons des procédures automatiques appelées algorithmes. Pour y pavenir nous devons:

1) définir P et ce qu'il contient
2) trouver $\phi := P \rightarrow \mu$, une ou plusieurs procédures capables de fournir μ
3) caracteriser ϕ et μ vis-à-vis d'objectifs assignés à l'un et l'autre de ces paramètres.

Dans la littérature de *market design* il existe de nombreuses procédures bien définies aux propriétés attrayantes. Mais la procédure n'est qu'une partie du travail de définition d'un mécanisme d'affectation des places en crèche. C'est un outil dont nous disposons pour résoudre le problème. Par ailleurs, si la question ne portait que sur la procédure ϕ, les chercheurs et ingénieurs en informatique pourraient proposer d'autres solutions.

Qu'est-ce qu'un marché, et pourquoi est-ce important Notre travail ne consiste pas à proposer une mise à jour des logiciels des directions petite enfance des territoires – même si cela peut passer par là – mais à concevoir des *places de marché*. Si nous utilisons ces termes, c'est parce que nous nous intéressons aux *transactions* que ce système permet d'organiser et à ce que la *société* tire d'une telle organisation. Les économistes[26] ont tendance à adopter des définitions assez larges de ce qui constitue une transaction et (donc) un marché (Satz 2010a). Des précisions s'imposent.

En bref, un marché est une institution qui facilite les transactions. Il peut être formel ou informel, conçu, encadré, physique, virtuel, utilisant un système de prix ou non… Un point important est que les caractéristiques du marché affectent qui négocie, ce qui est négocié et les termes de ces transactions (Li 2017a). Ainsi, les *places de marché* se présentent sous de nombreuses formes et ne sont pas uniquement, ni même principalement, des marchés de biens et services dont la seule fonction est la découverte des prix. Alvin A. E. Roth (2018) définit les places de

[24] De même que certaines recettes peuvent être préparée entièrement par des robots ménagers.

[25] Ce n'est pas tout à fait vrai dans cette expérimentation en particulier car nous introduisons dans la procédure ϕ un nombre aléatoire assigné à chaque dossier qui permet de départager les familles qui ont le même score de priorité et constitue l'élément le plus important pour la suite de la recherche. Néanmoins, ce tirage au sort n'est pas nécessaire au bon fonctionnement de ces procédures. Une fois le tirage réalisé, le problème est cette fois entièrement déterministe.

[26] dont nous faisons partie.

marché comme[27] “*l’infrastructure, les règles et les coutumes par lesquelles les informations sont échangées et les transactions effectuées, qui peuvent être des parties relativement petites de grands marchés. Les participants peuvent avoir de vastes ensembles de stratégies, c’est-à-dire de nombreuses options à leur disposition au-delà de celles disponibles sur une place de marché particulière*”.

Sur de nombreux marchés, nous nous soucions de savoir avec qui nous échangeons. Dans ces configurations, nous ne pouvons pas nous contenter de choisir ce que nous voulons, même si nous en avons les moyens : nous devons aussi être choisis. Le marché des modes d’accueils fait partie de ces marchés d’appariement (*matching markets*). Il existe des lieux d’inscriptions centralisés (les guichets uniques) donnant accès à une place de marché : les commissions d’admission des modes d’accueil (CAMA). Ces dernières sont en charge de l’appariement entre les offres de différents gestionnaires et les demandes des parents. Bien qu’elles proposent principalement des places en établissements d’accueil collectif du jeune enfant (EAJE) gérés, de près ou de loin, par le territoire, certaines proposent aussi des places dans des crèches associatives, des crèches familiales, des crèches d’entreprises, hospitalières ou universitaires,… En outre, cette place de marché n’est pas l’unique lieu des transactions pour les modes d’accueils. Les parents ont également d’autres options. Ils peuvent passer un contrat privé avec une assistante maternelle ou une crèche privée, s’occuper eux-mêmes de leurs enfants, définir une organisation repartissant la garde entre eux, des membres de leur famille ou d’autres liens sociaux, etc.

V.2 Principal résultat théorique

L’enjeu crucial des choix des décideurs Pour un marché donné et une définition de la stabilité, une appariement stable peut ne pas exister, ou il peut y en avoir plusieurs. Il existe modèles, algorithmes et façon de poser le problème qui peuvent produire des *résultats* très différents, c’est-à-dire en termes de distribution des affectations, ainsi qu’un ensemble de *conséquences*. Comme le souligne Li (2017a), une conséquence est plus riche qu’un résultat, c’est une “*description des effets du marché sur le monde*”. Il s’agit alors de peser et de résoudre les compromis entre différents objectifs et contraintes. Les modèles et la théorie aident à clarifier quels sont (certains de) ces compromis.

D’expérience, de nombreuses institutions locales collectent les préférences sur les jours de la semaine et attribuent les dossiers en conséquence[28]. Ce simple changement dans la manière dont les préférences sont collectées a des conséquences importantes sur la définition de la stabilité.

Les décideurs sont familiers avec les détails de leur environnement, et pourtant ils ne savent souvent pas comment formuler leurs objectifs en termes précis ou réalisent qu’ils sont conflictuels. Cependant, la théorie aide à fournir “*des orientations sans prescrire une théorie éthique entière*” (Li 2017a).

De nombreux objectifs politiques dépendent de ce qui est réalisable. Par exemple, un décideur politique pourrait penser qu’il est moralement obligatoire de mettre en place un système qui est optimal au sens de Pareto[29] pour les familles et respecte les priorités. Mais cela n’est pas possible (Abdulkadiroğlu and Sönmez 2003) donc cela ne peut pas être une contrainte morale ou politique. Lorsque l’attribution se fait en allouant des places sur les jours de la semaine, les développements récents sur l’appariement avec des contraintes multidimensionnelles pour la garde d’enfants (Kamada and Kojima 2023) et les relactions de réfugiés (Delacrétaz, Kominers, and Teytelboym 2023) fournissent différentes définitions de la stabilité et des algorithmes pour trouver des affectations stables avec d’autres propriétés souhaitables. Un résultat important de ces articles est qu’aucun algorithme ne peut respecter les priorités[30] et ne pas laisser de places vides. Encore une fois, les deux ne peuvent pas être réalisés, donc cela ne peut pas être un objectif politique. Cette littérature récente fournit des solutions théoriques bien adaptées au problème d’attribution de places en crèche avec des contraintes multidimensionnelles.

[27] “*infrastructure, rules, and customs through which information is exchanged and transactions are made [that] can be relatively small parts of large markets. Participants may have large strategy sets, i.e., many options available to them beyond those available in any particular marketplace*”

[28] En l’absence de jours spécifiques, un enfant occupe un siège, qui est un bien indivisible et unitaire. Cependant, dans des contextes avec des jours spécifiques, un enfant peut prendre certains jours et le reste peut être attribué à un autre avec des préférences complémentaires, ou non. Une place n’est plus indivisible et unitaire.

[29] Dans ce contexte, une affectation Pareto efficace implique que la distribution des places en crèche ne peut pas être améliorée d’une manière qui bénéficie à au moins un participant sans nuire à un autre participant.

[30] avec différentes définitions dans les deux articles.

Cependant, la mise en œuvre dans la vie réelle ajoute une autre couche de complexité. En effet, une autre caractéristique importante du système français est l'existence de contraintes de diversité dans toutes les structure et les procédures d'attribution. Les contraintes de diversité impliquent que les capacités au sein des crèches sont divisées en *groupes*, c'est-à-dire en ensembles de capacités avec des règles de priorité attachées aux *groupes*. Par exemple, certaines capacités seront réservées aux enfants de 6 à 12 mois uniquement, et d'autres aux enfants de 12 à 24 mois. Ces contraintes de diversité peuvent être *rigides* lorsque les groupes n'acceptent qu'un seul groupe, comme dans les exemples précédents. Parfois, les contraintes sont *flexibles* et définissent un ordre de préférence sur les groupes. Par exemple, un groupe peut accepter des enfants de 12 à 24 mois mais aussi des enfants plus âgés s'il y a suffisamment de capacité pour les accueillir. Alors que les formes les plus courantes de contraintes de diversité sont les groupes d'âge, les décideurs politiques et/ou les gestionnaires définissent de nombreuses autres formes de contraintes de diversité[31].

Trouver l'affectation préférée des familles sous un ensemble de contraintes Ces contraintes de diversité jouent un rôle important dans le manque de transparence des procédures d'attribution de places en crèche. En effet, les commissions d'attribution ne trient pas simplement les dossiers par priorités dans une crèche, elles considèrent l'attribution au sein de groupes, dont les définitions varient à l'intérieur et entre les crèches, générant ainsi plus ou moins de concurrence entre les groupes et donc des variations dans les probabilités d'attribution. Dans un appariement donné, il peut y avoir des enfants avec des scores de priorité bas attribués à une crèche que les parents ayant une priorité plus élevée voulaient mais ont été rejetés. Ainsi, la notion de stabilité ne peut pas être uniquement basée sur les scores de priorité, elle doit prendre en compte la répartition des capacités en groupes et les priorités à l'intérieur et entre les groupes.

Fournir une notion de stabilité dans un cas avec des contraintes multidimensionnelles et de diversité est donc à la fois pertinent sur le plan théorique et nécessaire pour assurer la transparence et le respect du processus.

Notr résultat principal pose la définition d'une places de marché pour l'attribution des places en crèche (DAM) comme trois éléments que les décideurs doivent choisir :

1) Une version du problème : qu'ils considèrent les demandes avec des jours spécifiques ou non ;
2) Une partition des capacités en *groupes*, c'est-à-dire un ensemble de capacités avec des règles de priorité attachées dans chaque crèche pour définir les contraintes de diversité ;
3) Une définition de *l'équité* : avec des scores de priorité et si les mécanismes doivent éliminer l'envie justifiée, tolérer certaines petites déviations ou considérer uniquement les demandes initialement réalisables.

En nous appuyant sur le travail de Ehlers et al. (2014) dans un cas sans attribution de jours spécifiques, Kamada and Kojima (2023) et Delacrétaz, Kominers, and Teytelboym (2023) dans un cas avec attribution de jours spécifiques, nous proposons des notions de stabilité pour différentes versions du problème d'attribution de places en crèche avec des contraintes de diversité.

Notre principal résultat – présenté dans le Théorème 2 – énonce ensuite que pour chaque DAM, nous pouvons trouver l'attribution équitable optimale unique pour les familles (SOFA), qui respecte la définition *choisie* de l'absence d'envie. Nous fournissons également une définition supplémentaire de stabilité sur les *demandes initialement réalisables*. À partir d'un SOFA avec toutes les demandes, nous montrons dans le Théorème 1 que supprimer les demandes initialement irréalisables est préférable pour tous les parents et donc, améliore au sens de Pareto l'attribution qu'ils reçoivent. Dit autrement, les personnes avec des scores élevées dont le voeu ne peut être réalisé dès le départ en raison d'un manque de capacité génèrent des externalités négatives sur les dossiers moins prioritaires qui pourraient être affectés. En retirant ces voeux, les personnes à qui on retire les voeux ne sont pas pénalisés puisque ces voeux ne pouvaient de toute façon être satisfaits ; mais les retirer permet à d'autres familles d'avoir parfois un meilleur choix. En pratique, ce théorème s'avère le plus utile dans les seconds tours où il y a beaucoup plus de jours vides dans les groupes.

[31] par exemple, pour les parents dont les horaires varient au fil des mois ou avec des horaires de travail décalés, ceux qui sont dans une politique active du marché du travail etc.

Les avantages des outils développés Une fois que les décideurs politiques définissent leur DAM, les procédures proposées offrent plusieurs avantages distincts par rapport aux pratiques actuelles:

1) Ces outils informatiques peuvent rapidement traiter un grand nombre de demandes, ce qui permet de gagner du temps pour ceux impliqués dans l'organisation et la participation aux commissions d'attribution.
2) Nos modèles clarifient ce qui peut être fait et ce qui ne peut pas l'être, ainsi que les compromis de chaque mécanisme d'attribution. D'un point de vue normatif, ils garantissent des affectations *justes* au sens où ils éliminent les *envies justifiées* sur la base d'une définition claire et choisie. Les décideurs politiques choisissent s'ils veulent respecter strictement les priorités ou accueillir davantage d'enfants en autorisant une adaptation faible et/ou en ne tenant compte que des demandes initialement réalisables.
3) Toutes les décisions d'attribution sont traçables et peuvent être expliquées à chaque famille. Pour chaque attribution, nous fournissons un tableau pour chaque crèche qui justifie chaque décision. Une transparence totale est possible, si les décideurs politiques souhaitent aller vers cette direction.
4) En garantissant que les priorités sont toujours respectées et individuellement justifiables, ils peuvent créer un bien public: ces mécanismets assurent une forme de *justice procédurale* et peuvent déplacer les préoccupations éthiques et de justice vers la *justice distributive*. Puisque les priorités sont toujours respectées, la société peut débattre de ce qui constitue des inégalités justes.
5) Les décideurs politiques peuvent définir des distributions d'affectation cibles en choisissant les parts de places qu'ils veulent attribuer à certains groupes. Nos modèles sont suffisamment flexibles pour prendre en compte de nombreuses contraintes. En particulier, la définition de quotas de diversité (flexibles) est un outil puissant pour garantir que l'affectation satisfait certaines exigences de distribution. Sans notre procédure, ce type d'objectifs politiques est très difficile à atteindre. Ou du moins, difficile à justifier.
6) Ces outils peuvent également être utilisés pour simuler les effets de la modification de certains paramètres et informer les décideurs politiques en utilisant des scénarios *et si*.

V.3 Contribution des procédures d'affectations à l'effet Matthieu

La deuxième partie de notre recherche se concentre sur l'application pratique de ces mécanismes automatisés d'attribution de places en crèche. Nous utilisons principalement des données provenant d'un grand centre urbain pour lequel nous disposons de quatre années de fonctionnement. Il est important de noter que les ensembles de données sont suffisamment riches pour une compréhension globale des critères, des poids et de la formule du score de priorité, et englobent une réforme de ce dernier au milieu de la période étudiée. Ce mécanisme est basé sur les demandes spécifiques aux jours de la semaine avec des contraintes flexibles sur les groupes d'âge. La deuxième étude de cas ne concerne qu'une seule attribution basée sur la version du problème de choix d'école.

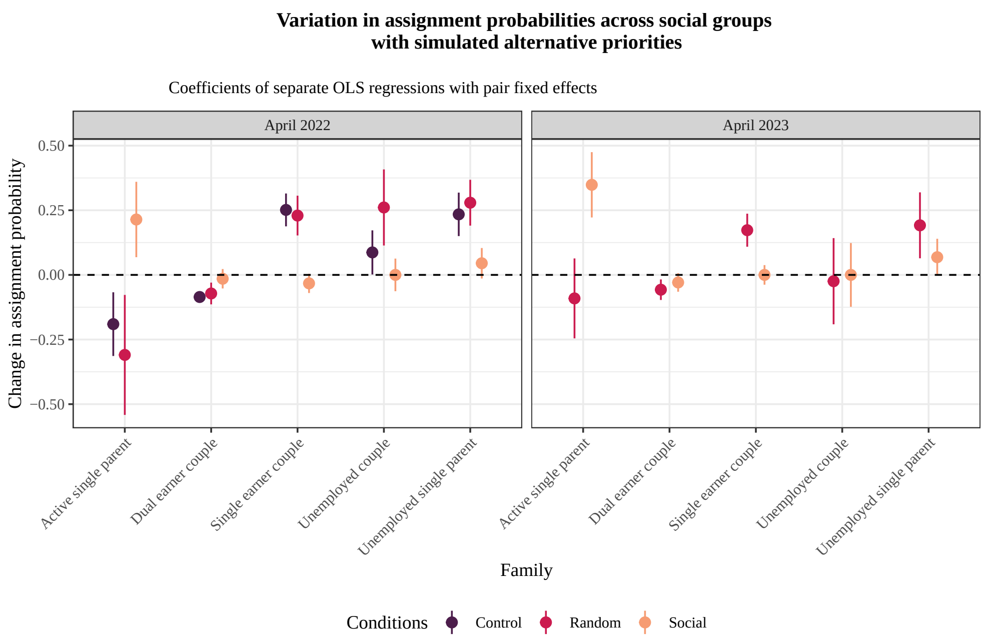
Nos conclusions reposent sur quatre évaluations empiriques clés fortement inspirées de l'audit en quatre étapes de l'algorithme proposé par Kasy and Abebe (2021). Premièrement, nous examinons les propriétés de nos algorithmes et comparons les résultats de différents mécanismes avec l'attribution finale. Deuxièmement, nous exploitons une réforme des priorités au milieu de la période de l'étude de cas I pour mesurer ses effets et ceux de règles de priorité alternatives sur les inégalités et la ségrégation entre les groupes sociaux. Troisièmement, nous étudions l'impact de l'inscription stratégique et ses conséquences sur les inégalités par mois de naissance. Quatrièmement, nous utilisons les données de l'étude de cas II pour démontrer le rôle critique des éléments et de la distribution des groupes.

Le rôle central des priorités et des stratégies des parents Dans la première étude de cas, le score de priorité est composé de quatre éléments : i) un désavantage pour les non-résidents, ii) des pondérations sociales sur les groupes basées sur la distribution conjointe de l'emploi et du statut relationnel, iii) Des poids de « circonstances » (concernant 15% des dossiers) pour prioriser des familles avec des situations plus compliquées (handicap, mobilité professionnelle, accompagnment par l'aide sociale à l'enfance etc.), et iv) le temps écoulé depuis l'inscription, ouverte à partir de la déclaration de grossesse.

Ce dernier élément encourage les inscriptions précoces et explique la plupart des variations de priorités. Le groupe majoritaire - les couples bi-actifs - ont souvent besoin d'un mode d'accueil tôt après la naissance et s'inscrivent de manière stratégique dès que possible. En utilisant la simulation d'une affectation contrefactuelle avec des priorités alternatives, nous avons montré que la pondération des groupes sociaux peut efficacement modifier les probabilités d'affectation. Cependant, permettre aux parents de manipuler leurs priorités grâce à une inscription précoce pénalise fortement les familles monoparentales.

Nous l'illustrons en reproduisant une des figures du Chapitre 1 dans la Figure 6. Cette dernière compare les probabilité d'admissions entre groupes sociaux lors de la commission du printemps 2022 et 2023 en simulant des priorités différentes. En particulier, nous simulons une affectation ne tenant compte que des préférences des familles et contraintes d'âges, départageant au hasard les dossiers pour le reste, et une affectation avec l'ensemble des critères excepté l'ancienneté de la demande. La première montre que les poids donné aux différents groupes permet bien d'augmenter les priorités des groupes favorisés. La seconde montre que les seuls groupes pour qui la probabilité moyenne change lorsqu'on ajoute le critère d'ancienneté sont les familles monoparentales. Surtout celles actives qui ont le plus haut poids social mais qui, du fait d'inscription généralement plus tardives, se retrouvent largement évincées.

Figure 6: Le critère d'ancienneté évince les familles monoparentales



Sources: ISAJE, Case Study I − 2022 : 2023 − first rounds only.
Notes: We simulated assignments using alternative definition of priority scores while all other parameters are held constant.
We stacked simulations in a database and present the coefficients of separate regressions by social groups of a dummy for assignment over scenario dummies and individual fixed effects. Confidence intervals are clustered at the individual level.
'Actual' assignment is the reference, 'Control' is the scenario without the reform i.e. the 2021 weights,
'Random' is the scenario without priorities i.e. students are sorted by decreasing lottery realisation with age buckets,
and 'social' use the priorities without weights for time since registration. In 2022, it uses reformed weights.
We use cluster robust standards errors at the individual level to build 95% confidence intervals.

Cela s'explique en partie parce qu'une inscription précoce ne peut pas être une stratégie pour ceux qui sont devenus familles monoparentales plus tard dans leur grossesse. Ces inégalités sont aggravées par des pondérations sociales élevées pour les couples où un seul parent travaille, qui peuvent anticiper davantage, attendre plus longtemps et à moindre coût. Au cours des quatre années, les familles monoparentales et les couples où un seul parent travaille n'ont fait qu'observer une détérioration de leur situation. Le moment stratégique de l'inscription n'a pas le même rendement lorsque les grossesses sont déclarées plus tard dans l'année et crée des inégalités d'opportunités même parmi les parents stratégiques grâce à de grandes variations dans les probabilités d'affectation selon la date de naissance. Dans la deuxième étude de cas, la répartition des capacités entre les groupes d'âge est très déséquilibrée et loin de la demande correspondante. La forte discontinuité dans les probabilités d'affectation selon la date de naissance est une autre illustration des inégalités d'opportunités découlant de la structure des groupes.

Expliquer l'effet Matthieu par les choix des décideurs Les familles monoparentales occupent une place centrale dans le discours sur l'investissement social. Les politiques visant à concilier travail et vie familiale, telles que les arrangements formels de garde d'enfants, ont pour objectif à la fois de permettre aux parents de travailler et de contribuer au développement sain des enfants, réduisant ainsi les inégalités sociales dès l'enfance. Nous avons montré que dans ces deux contextes, les décideurs politiques influencent fortement la distribution des affectations et favorisent les ménages plus aisés et une entrée précoce dans la garde d'enfants. Le temps écoulé depuis l'inscription affecte négativement les familles mono parentales et les parents sans emploi. Le système de priorité actuel avantage fortement les familles stables : les couples stables, avec des emplois stables, qui connaissaient leurs besoins en matière de garde d'enfants et le moment de s'inscrire, très probablement avant la grossesse.

Nos résultats montrent que dans les cas étudiés ici, l'effet Matthieu est forcement accentué par des choix de conception et notamment la définition des critères de priorité. Si l'on pense aux objectifs ambitieux des politiques d'accueil du jeune enfants affichés par les politiques, ce type de distribution est très régressif et va à l'encontre d'objectifs d'investissement social (van Huizen and Plantenga 2018; Schmutz 2024). Ces résultats ont de fortes implications économiques et d'équité. D'un point de vue économique, l'effet Matthieu crée d'importantes pertes d'efficacité en réduisant les opportunités de développement cruciales pour les familles à faible revenu. Une autre façon de penser à ce problème est de considérer le surplus généré par l'appariement et les coûts d'opportunité que de telles priorités génèrent pour les parents, les décideurs politiques et la société. Une première chose importante à noter est que les poids actuels génèrent des avantages privés réels, tandis que les coûts d'opportunité sont principalement incertains et à long terme. Neimanns (2022) soutiennent qu'il existe des gains électoraux à court terme à favoriser les familles plus aisées et, en pratique, c'est bien ce que nous observons. De plus, les frais augmentent avec les revenus, il est donc également moins cher de les favoriser. Les familles aisées bénéficient également de la garde collective qui est en général moins chère que les assistantes maternelles ou les services de garde privés. Leur carrière peut également être moins affectée par la pénalité liée à l'enfant, ainsi que le montrent la Figure 3 présenté plus tôt.

Les enjeux politiques de nos travaux Une ambiguïté entre le *market design* et la recherche en informatique réside dans la façon dont les utilisateurs ou les décideurs politiques perçoivent le rôle de tels mécanismes. Dans notre travail sur le terrain, certaines municipalités étaient intéressées par nos algorithme car ils réduisent la charge de travail en coulisses des commissions et les considèrent d'abord comme un système informatique performant. D'autres au contraire, valorisaient les propriétés d'équité et de transparence de nos outils; certaines allaient même jusqu'à communiquer les tirages au sort aux parents pour justifier les affectations.

En général cependant, l'automatisation s'est accompagnée d'un minimum d'informations fournies aux familles, principalement pour préserver l'intégrité du protocole de recherche et éviter les réactions politiques négatives. Comme nous l'avons vu, les mécanismes sont adaptés aux DAM basés sur nos modèles d'agents économiques. Nous comprenons qu'ils peuvent réagir à différents designs ou fonctionnalités. Le market design est fortement lié à la théorie des jeux. Un jeu est basé sur des règles et ces règles définissent qui joue au jeu et quelles stratégies chaque joueur peut mener. L'adoption officielle de nos outils suscitera probablement différentes réactions de la part des familles, des professionnels de la petite enfance et des décideurs politiques.

Les dynamiques de pouvoir locales sont importantes et les élus apprécient leur autonomie. Même la mission interministérielle confiée à un membre élu de l'Association des maires de France (AMF) n'a pas pu formuler de recommandation sur les critères qui pourraient/devraient être utilisés pour les priorités. L'opposition de l'AMF à ce projet a d'ailleurs été clairement exprimée, comme discuté dans le Chapitre 1. Notre approche a notamment été perçue à la fois comme une menace pour l'autonomie des territoires et a été amalgamée avec des outils informatiques. L'aspect politique de cette recherche n'est pas secondaire. Nous avons modifié les structures de marché. Nous sommes comptables des affectations réalisées ; ces affectations ont déterminées quelles familles ont reçu une place et celles qui n'en ont pas reçu. Pour respecter l'éthique et l'intégrité scientifique, notre travail a été discuté et approuvé par un comité scientifique international composé de chercheurs de différents domaines (sociologie, économie, sciences de l'éducation, etc.). En tant que chercheurs, nous devons reconnaître notre responsabilité et une partie de celle-ci consiste à rendre compte de ce que nous avons appris. Bien que nous ayons rencontré de la résistance au stade initial, nous avons reçu de nombreux retours positifs de la part de ceux avec qui nous avons travaillé, tels que des confirmations informelles de réductions d'interventions de la part des élus ou des plaintes de la part des familles.

La recherche ISAJE est à ma connaissance une première mondiale. Outre la réflexion sur le développement des outils permettant un meilleur fonctionnement des guichets uniques, nous avons produit des données inédites dont nous débutons seulement le traitement. Entre 2020 et 2023, nous avons pu convaincre neuf grandes agglomérations urbaines de participer à cette expérience, acceptant d'automatiser leur procédure d'attribution de places en crèche en utilisant des tirages aléatoires pour classer les familles ayant le même niveau de priorité. Une collectivité est restée impliquée pendant quatre ans, deux autres pendant trois ans, trois autres pendant deux ans, et trois autres pendant un an. Au total, environ 20 000 familles ont été soumises à nos procédures automatisées pour 5 000 places en crèche. Après chaque affectation, nous avons simulé un million d'attributions, en modifiant uniquement le tirage aléatoire pour calculer les probabilités individuelle d'être acceptée, d'avoir son "premier choix", etc. Ce projet est toujours en cours en 2024.

VI Contribution du Chapitre 2

Dans le chapitre 2, j'analyse les effets sur la participation au marché du travail et la pauvreté du programme *Reliance* introduit dans la Section IV.

Le programme *Reliance* est profondément enraciné dans le paradigme de l'investissement social par sa cible, ses objectifs politiques et sa conception. Il mobilise des ressources importantes de diverses institutions, notamment la Cnaf et la Caf du département, la Caisse des dépôts, ainsi que le conseil départemental. Il *investit* environ 2 800 euros par participant - quatre fois le montant habituel des dépenses par bénéficiaire du RSA - dans le but de favoriser la participation au marché du travail et, finalement, de réduire la pauvreté des familles très vulnérables.

Dans la littérature, ces éléments sont généralement prédictifs d'effets plus importants sur la participation au marché du travail. Par exemple, Bloom, Hill, and Riccio (2003) utilisent des données de trois expériences d'assignation aléatoire à grande échelle, menées dans plusieurs sites aux États-Unis, et montrent que l'accent mis sur un retour rapide à l'emploi, un soutien personnalisé, une utilisation limitée de l'éducation de base et des effectifs réduits augmentent la taille de l'effet sur l'emploi et les revenus. Cependant, la revue systématique de Gorey (2009) montre que les effets positifs ne sont observés que lorsque le programme offre un accès à des services de garde abordables et diminuent lorsque les taux de chômage augmentent et que les emplois deviennent plus difficiles à trouver. En revanche, la revue des effets des mesures actives du marché du travail pour les femmes en Europe, menée par Bergemann and Van Den Berg (2008), fait état principalement d'effets positifs, plus prononcés pour les femmes que pour les hommes, en particulier pour les programmes de formation. Cependant, la grande majorité de ces recherches (39) sont des quasi-expériences, tandis qu'ils ne recensent que 4 essais randomisés donnant des résultats opposés (deux positifs, deux négatifs).

Le principal défi de cette recherche est de garantir une évaluation qui informe véritablement les politiques publiques. L'analyse est donc guidée par les questions causales suivantes :

> *Le programme augmente-t-il la participation au marché du travail ? Réduit-il la pauvreté ?*

Les données exploitées portent sur 844 parents isolés encouragés des quatres premières cohortes ayant été invités à des réunions présentant le programme, avec une moyenne de 38% d'inscription, et 828 dans le groupe témoin n'ayant reçu aucune intervention. À l'instar de la méthodogie présentée dans la Section IV, j'utilise des méthodes de variable instrumentale pour identifier l'effet moyen du programme sur les participantes.

VI.1 Rare expérimentation aléatoire en France

Il existe un manque criant d'évaluation d'impact de qualité mesurant les effets des programmes d'aide sociale à l'emploi en France (P.-H. Bono et al. 2021). Dans une revue de l'effet du soutien social sur diverses dimensions, Cervera et al. (2017) notent seulement quelques analyses quantitatives en France. Les exceptions se concentrent principalement sur les résultats en matière d'emploi pour les chômeurs ou les jeunes adultes. Une revue systématique de la littérature sur l'effet des mesures actives du marché du travail sur les chômeurs de longue durée entre 2000 et 2015, menée par Abadia et al. (2017), ne signale qu'une seule étude impliquant des bénéficiaires de l'ancien dispositif de revenu minimum (RMI), avant l'introduction des prestations en emploi (Crepon et al. 2013).

Cette étude présente une évaluation expérimentale unique et contribue à combler ce vide dans la littérature sur les politiques d'aide sociale à l'emploi pour les familles monoparentales en situation de pauvreté. Bien qu'il soit toujours du devoir d'un chercheur de fournir des preuves claires et de haute qualité avec une méthode empirique transparente, les enjeux de cette étude semblent notablement plus élevés. Pour garantir l'intégrité de la recherche, toutes les analyses et les résultats ont été préenregistrés, ce qui permet de réduire les pressions politiques potentielles et d'éviter la sélection arbitraire des résultats. De plus, ces recherches sont menées avec des outils informatiques permettant de répliquer les analyses à l'identique dans un souci de transparence et de science ouverte.

Le design précis se base sur un plan d'encouragement randomisé par blocs, c'est-à-dire qu'à partir de l'échantillon initial, j'ai d'abord construit des groupes de familles homogènes selon le nombre d'années passé au RSA (2 à 5 ans, 5 à 10 ans ou plus de 10 ans), insciption à Pôle emploi et nombre d'enfant (1, 2, 3+). Le tirage au sort affectant les familles au groupe témoins et encouragé et réalisé au sein de ces groupes construit à partir de l'interaction de ces caractéristiques. Cette étape sert d'une part à garantir un équilibre stricte - et non seulement en moyenne - entre les bras expérimentaux et à gagner en précision. En effet, ces variables ont été choisies pour leur possible effet sur la participation et les variables de résultats. Plus celles-ci sont prédictives, plus les estimateurs de taille d'effet sont précis. Enfin, l'équilibre entre les groupes permet de faire des analyses d'hétérogénéité en comparant ces sous-groupes.

VI.2 Absence d'effet moyen du programme sur l'emploi et la pauvreté

Les résultats peuvent être résumé ainsi:

1) Le taux de participation est en moyenne de 38 %, passant de 28 % à 48 % de la première à la quatrième cohorte. Cette amélioration découle très probablement des ajustements apportés au processus de recrutement, notamment i) des invitations plus menaçantes ii) des séances d'information déplacées sur les lieux du programme iii) le passage de réunions collectives à des entretiens individuels en face-à-face avec les responsables du projet et iv) des témoignages d'anciens participants.
2) Les personnes qui adhèrent au programme ont plus de chances d'être dans la trentaine, parmi les plus pauvres, d'avoir moins qu'un diplôme d'études secondaires et d'être inscrites à l'Agence pour l'emploi.
3) Le programme ralentit le taux de placement dans les emplois pendant sa première moitié, ce qui entraîne un fort effet de verrouillage se traduisant par un taux de pauvreté plus élevé, un revenu disponible et un emploi. Ces effets négatifs anticipés s'estompent à la fin du programme et il n'y a pas d'effet moyen dans la période post-traitement.
4) Le programme augmente progressivement le montant des transferts monétaires reçus, bien que cet effet soit entièrement médiatisé par l'augmentation de la taille de la famille.
5) L'analyse de l'hétérogénéité des effets du traitement révèle des schémas troublants, notamment parmi les revenus élevés/bas et le nombre d'enfants au début de l'étude, où les effets sur les revenus disponibles et la participation au marché du travail présentent des tendances divergentes, suggérant des changements dans la composition des revenus, corroborés par l'augmentation des transferts monétaires à la fin de la période.

Les effets de ce programme sur la participation au marché du travail et le revenu disponible des mamans-solos pauvres en France sont très négatifs. Il n'a pas réussi à augmenter l'emploi ou à sortir ces mamans-solos de la pauvreté.

VI.3 L'importance du biais de sélection

Dans cette expérience, les participants semblent se diviser en deux groupes : les plus pauvres avec un niveau de diplôme inférieur au bac d'une part, et celles plus proches du marché du travail d'autre part. Pour les premiers, le programme réduit l'emploi mais n'a aucun effet sur les revenus disponibles, tandis que pour les seconds, il n'a aucun effet sur l'emploi mais réduit significativement leurs revenus. Cela signifie uniquement que les plus pauvres ont trouvé d'autres sources de revenus tandis que ceux qui sont plus proches du marché du travail gagnent moins qu'ils n'auraient dû s'ils n'avaient pas participé.

Le programme attire celles ayant le plus haut niveau d'emploi potentiel, mais n'augmente pas la participation au marché du travail. Le biais de sélection est si fort que les estimations utilisant la prochaine meilleure stratégie d'identification mise en œuvre avec des estimations robustes modernes ne parviennent pas à inclure les résultats expérimentaux dans les intervalles de confiance. Sans affectation aléatoire, j'aurais sans doute conclu à tort que le programme augmente l'emploi.

La Figure 7 reproduit l'un des principaux résultat du Chapitre 2 illustrant ce point. Dans la partie haute, je représente la moyenne et l'intervalle de confiance à 95% du taux d'emploi chaque mois autour du tirage au sort, en fonction du groupe d'assignation et de participation. Le groupe encouragé est donc divisé entre les participantes (*treated compliers*) et les jamais partantes (*Never takers*).

Ce qu'il faut noter, c'est que les participantes sont nettement plus en emploi à l'issue du programme que ne le sont celles du groupe témoins et les jamais partantes. Avant le tirage au sort, les trajectoire étaient parfaitement confondues ce qui semble indiquer qu'avant l'intervention, les groupes étaient sur des trajectoires d'emploi parallèles. Dans ces conditions, une stratégie en double différences semble indiquée et les estimations présentées dans la partie basse de la figure reporte ces estimations.

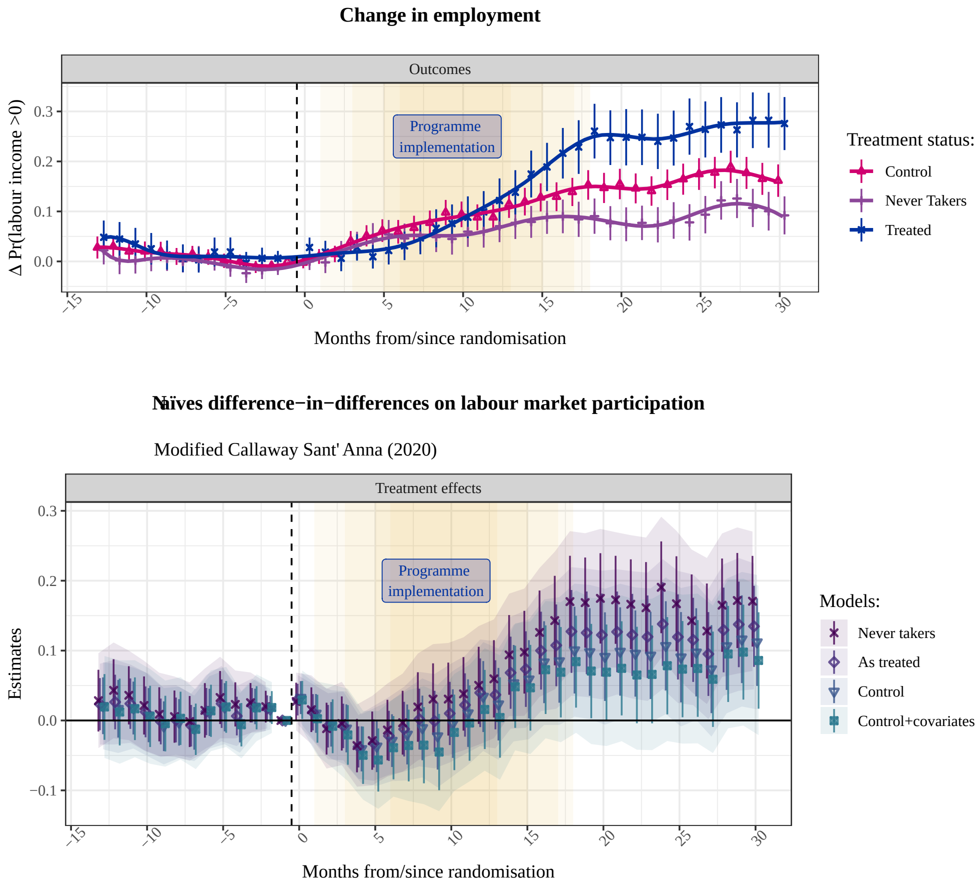
Mais les progrès observervés pour les participantes ne reflètent pas l'effet du programme mais bien la sélection *endogène* de maman solos plus dynamiques et mobilisées par l'intervention. Intuitivement, le groupe de contrôle contient, comme le groupe encouragé, des jamais participantes d'un côté et des mobilisables de l'autre. Les mobilisables sont celles qui auraient participé si le programme leur avait été proposé. Si nous pouvions observé qui sont les mobilisables dans le groupe témoins nous pourrions les comparer avec les participantes car il s'agit du contrefactuel pertinent. Cependant, cette caractéristique n'est pas observable, il s'agit d'une participation potentielle, latente.

Les doubles différences utilisent des caractéristiques observables et un groupe de contrôle choisi en supposant qu'il correspont au groupe de mobilisable, mais ce n'est pas le cas. Ainsi, les doubles différences intégrants les jamais-partantes présentent des résultats beaucoup plus élevées car en l'absence du programme, leur trajectoire aurait été plus basse que celles des participantes si le programme n'existait pas.

Coûts d'opportunité des fonds publics Dans cette expérience, les décideurs ont révélé leur volonté de sortir ces familles de la pauvreté, en dépensant environ 2800 € par participante. En se fixant un objectif d'emploi de 10 pp pour le succès, ils étaient prêts à soutenir 10 familles monoparentales pour qu'une seule trouve un emploi en moyenne, acceptant implicitement un coût de 28 000 € par emploi attendu. Cela dépasse le coût total pour un employeur d'un emploi à plein temps d'un an au salaire minimum (qui est d'environ 22 000 € par an) tout en reposant sur des effets très incertains. En réalité, le programme n'a pas augmenté l'emploi et a plutôt augmenté les transferts sociaux. Cela fait écho à l'un des arguments motivant le projet "*Territoire Zéro Chômeur*", où les chômeurs de longue durée se voient offrir des contrats permanents au salaire minimum à temps plein.

Une fois encore, cette expérience à enjeux élevés n'a pas randomisée mais le service statistique du Ministère du Travail (DARES) a utilisé le couplage pour évaluer les effets et a trouvé des effets positifs significatifs sur l'emploi et le bien-être des participants (DARES 2021). En revanche, Kasy and Lehner (2023) évaluent une politique similaire en Allemagne et utilisent un plan de recherche très rigoureux. Ils examinent l'effet moyen de traitement individuel en utilisant un plan d'appariement aléatoire, évaluent l'effet agrégé en utilisant un contrôle synthétique préenregistré au niveau de la commune et une comparaison avec les individus dans les communes témoins. Cela permet d'identifier les retombées. Ils constatent des impacts positifs de la participation au programme sur le bien-être économique et non économique, mais pas sur la santé physique ou les préférences. Au niveau de la commune, ils constatent une forte réduction du chômage de longue durée et aucune retombée négative sur l'emploi.

Figure 7: Effets de sélection et résultats en double-différences



Sources: ALLSTAT 2017−01−01 to 2023−06−01 cohorts 2018 to 2021.
Notes: The dependent variable is the long difference between employment at any relative month and the month before randomisation.
Top panel:
− Points indicate simple means over cohorts 2018 to 2021 in relative time since randomisation with 95% error bars by treatment status.
− Lines are conditional means estimated with spline regressions.
Bottom panel:
Separate difference−in−differences for each cohort using the doubly robust estimator proposed by Sant' Anna and Zhao (2020) aggregated following Callaway−Sant' Anna (2020), i.e. separate DiD for each cohort and weighted average of cohort x time treatment effects.
Covariates are measured at the month before random assignment and include baseline level, number of years receiving RSA, number of children and unemployment registration status (uninteracted blocking variables).
French citizenship, High/Low education, favourable assessment, receiving each social transfers, child support, children between 3 to 5 and at least one child over 16, quartiles of age, income per capita, taxable income.
The error bars indicate the 95% confidence intervals based on cluster−robust standard error adjusted at the household level. The shaded areas represent the FWER adjusted 95% confidence levels estimated using wild cluster bootstrap.

Il semble important de souligner le risque pris en poursuivant des politiques d'activation et de remettre en question les motivations et justifications de ces choix, compte tenu des coûts impliqués, du risque d'échec, et du peu de considération pour les conséquences en cas d'échec. Une utilisation différente de ce même budget pourrait-elle atteindre les objectifs des décideurs politiques ? Pourrait-on faire mieux ? Par exemple, plusieurs études récentes soutiennent l'idée qu'un transfert monétaire significatif produit des effets positifs durables sur la pauvreté (Jones and Marinescu 2022) et l'éducation des enfants (Barr, Eggleston, and Smith 2022) dans des contextes très différents. Plus récemment, une expérience randomisée à grande échelle de transferts d'argent inconditionnels aux États-Unis a montré qu'elle augmente les dépenses des mères pour leurs enfants et le temps passé avec eux (Gennetian et al. 2022). Cependant, ces montants doivent être suffisants pour répondre aux besoins des familles, sinon de telles interventions n'ont pas d'effet durable (Jaroszewicz et al. 2022). Dans l'ensemble, les programmes d'assistance sociale dans les pays à revenu élevé sont insuffisants pour préserver la santé et le bien-être des populations socialement défavorisées, ce qui indique que la portée et la générosité des programmes existants sont insuffisantes pour compenser les conséquences négatives associées à la pauvreté (Shahidi et al. 2019).

VII Contribution du Chapitre 3

Ce chapitre, co-signé par Alexandra Galitzine, porte sur les incitations monétaires à la reprise d'emploi pour les familles monoparentales.

Le débat entourant les politiques de redistribution via une fiscalité progressive et des transferts généreux tourne autour de la tension entre justice sociale et efficacité économique. D'une part, la redistribution est essentielle pour assurer le bien-être économique des moins fortunés, reconnaissant que les différences de revenus découlent souvent de facteurs hors du contrôle des individus, tels que les capacités innées, l'origine sociale ou simplement la (mal)chance (Fleurbaey and Maniquet 2018). D'autre part, les mesures redistributives peuvent entraver les incitations au travail tant chez les riches que chez les bénéficiaires de transferts, potentiellement compromettant la productivité économique globale (Ebert 2005; Piketty and Saez 2013). Au cours des trois dernières décennies, les États providence traditionnels ont été soumis à la pression de la consolidation fiscale et se sont orientés vers des politiques actives du marché du travail (PAMT) (J. P. Martin 2015; Crépon and van den Berg 2016a). Ces réformes sociales ont pris différentes formes à travers le monde mais ont radicalement remodelé les institutions du marché du travail et les mécanismes d'assurance sociale avec des programmes de retour au travail, des prestations en emploi, des exonérations de cotisations sociales, des initiatives de formation, d'assistance à la recherche d'emploi et de suivi, et ainsi de suite.

L'adoption par la France du Revenu de Solidarité Active (RSA) en 2008 est l'un de ces exemples et les réformes ultérieures en 2016 et 2019 – séparant la prestation en emploi en Prime d'Activité (PA) et augmentant les montants plus haut dans la distribution des salaires – soulignent les efforts continus pour inciter au travail et augmenter les revenus disponibles des travailleurs à bas salaire (Gurgand and Margolis 2008; Bargain and Doorley 2011; Bargain and Vicard 2014; Sicsic 2019).

L'enjeu de la connaissance et de l'accès à l'information sur les programmes sociaux Cependant, ces politiques reposent sur plusieurs prérequis. Tout d'abord, les personnes doivent comprendre comment les avantages modifient leurs incitations au travail. Lorsqu'on comprend mal les programmes fiscaux ou d'aide sociale auxquels on est soumis, on a tendance à les percevoir de manière simplifiée et à sous-réagir aux incitations. Des travaux récents montrent que les ménages sont souvent mal informés et ont parfois du mal à comprendre les non-linéarités du système. Ils s'appuient alors sur un taux d'imposition "*mentalement linéarisé*" (Rees-Jones and Taubinsky 2020; Caldwell, Nelson, and Waldinger 2023). L'incertitude peut résulter de caractéristiques plus complexes du code fiscal, telles que les régions de montée en charge ou des règles liées à la situation conjugale, etc. En revanche, fournir des informations supplémentaires peut déclencher de grandes réactions aux marges intensives, induisant de grands effets de distribution (Raj Chetty, Friedman, and Saez 2013; Raj Chetty and Saez 2013; Kostøl and Myhre 2021). Cependant, les moyens par lesquels les informations sont fournies semblent tout aussi importants; de nombreuses expériences d'information seule et de nudges se sont révélées inefficaces (Linos et al. 2020; Nyman, Aggeborn, and Ahlskog 2023) ou ont donné des effets beaucoup plus faibles que l'aide humaine (Castell et al. 2022; Finkelstein and Notowidigdo 2019; Bergman et al. 2019).

Deuxièmement, un autre prérequis est qu'il devrait effectivement y avoir des incitations au travail. Cela peut sembler évident compte tenu du montant des fonds publics dépensés dans les politiques actives du marché du travail. Cependant, cela peut ne pas être le cas pour tout le monde, et il peut exister de grandes frictions entravant les ajustements. Au niveau macro-économique, de nombreux universitaires ont souligné le paradoxe qu'en dépit des augmentations massives des dépenses publiques dans les politiques actives du marché du travail, le chômage de longue durée et la pauvreté n'ont pas été réduits (Vandenbroucke and Vleminckx 2011a; Jaehrling, Kalina, and Mesaros 2015; Van Winkle and Struffolino 2018). Au niveau microéconomique, un grand nombre de recherches montre que les politiques actives du marché du travail ont peu d'effets sur les résultats du marché du travail des familles monoparentales et ont souvent des effets préjudiciables sur leur santé et leur bien-être (Cook 2012; Pega et al. 2013; Gibson et al. 2018), parfois étendus à leurs enfants (Løken, Lommerud, and Holm Reiso 2018).

Dans le déroulement de cette expérience, la réforme de la prime d'activité en janvier 2019 a été adoptée de manière inattendue au dernier trimestre de la formation de la première cohorte alors que la deuxième était recrutée. Elle a été suivie par une forte augmentation du nombre de bénéficiaires, provenant en grande partie de ménages précédemment inéligibles et situés plus haut dans la distribution des salaires, mais aussi de nouveaux inscrits (Dardier, Doan, and

Lhermet 2022). Cependant, les évaluations actuelles sont incapables de démêler la réponse à la marge extensive – *i.e.* l'effet sur la reprise d'emploi – de la réduction du non-recours[32].

Notre cadre se révèle très adapté pour mesurer les réactions des familles monoparentales au système socio-fiscal nouvellement réformé. Bien que nous ne puissions pas mesurer l'effet de la réforme en soi, car toutes les cohortes ont été exposées, nous pouvons mesurer si le programme a eu des effets cohérents avec les incitations. En effet, le programme **Reliance** fournissait directement des informations individualisées et détaillées dans le cadre d'un accompagnement global d'une durée d'un an, susceptible d'avoir réduit, davantage encore, les divers freins à l'emploi. Dans ces conditions, notre expérience révèle des élasticités de travail *quasi-sans friction* pour les familles monoparentales au RSA longue durée.

En quelque sorte, nous exploitons cette expérience randomisée pour découvrir des réactions plus structurelles au système socio-fiscal, que nous avons donc documenté. Ce chapitre est donc guidé par deux questions de recherche:

1) *Comment le système socio-fiscal s'ajuste-t-il aux revenus et à la composition familiale ?*
2) *Comment les familles monoparentales réagissent-elles aux incitations du système socio-fiscal ?*

Exploiter la randomisation au delà de *"What work"* Les expérimentations aléatoires sont souvent critiquées pour plusieurs raisons. Tout d'abord, elles peuvent être coûteuses et complexes à mettre en œuvre, ce qui limite leur faisabilité dans de nombreux contextes. Ici, le coût de l'évaluation est très faible ; toutes les données proviennent de source administrative et en dehors des coûts de coordination et mon salaire, produire ces analyses causales n'ont pas occasionné de coûts additionnels. L'évaluation qualitative et les enquêtes menées et jusqu'alors non exploitées sont coûteuses précisément car elles produisent de nouvelles données.

En outre, l'approche "what works" souvent associée à ces expériences apparaît souvent limitée, car les recherches ne se focalisent que sur la mesure de l'effet sans en analyser les mécanismes. Dans cette recherche, nous montrons qu'une expérimentation aléatoire fournit beaucoup plus qu'une simple analyse de ce qui fonctionne ou non.

Lorsqu'on dispose d'une variable instrumentale crédible, toute la distribution des résultats potentiels est identifiée (A. Abadie 2003). Avec elle, les caractéristiques des mobilisables et des jamais-partantes. Dans cette expérimentation, le fait qu'aucune personne du groupe de contrôle n'ai participé facilite encore l'identification de paramètres importants. En examinant les réactions des bénéficiaires à un programme dans un cadre contrôlé, nous pouvons mieux comprendre les mécanismes et les implications des politiques sociales, offrant ainsi des perspectives précieuses pour informer les décideurs politiques et les praticiens.

Ici, nous nous intéressons en particulier à la distribution des revenus du travail des participantes si elles n'avaient pas participé. Cette quantité par nature inobservable peut être estimé à partir des comparaisons entre les groupes tirés au sort et avec l'information sur la participation. L'intuition derrière cette méthode réside dans le fait que le groupe de contrôle comprend à la fois des personnes qui n'auraient jamais participé au programme et des individus qui auraient pu être mobilisés (les mobilisables). En connaissant le poids de chaque groupe, il devient possible de réévaluer la distribution inobservée en repondérant les groupes. On parle de *poids causaux* pour cette méthode.

Notre article s'inspire principalement des rares exemples dans la littérature sur les effets de distribution des politiques d'activation. En particulier, notre travail est le plus proche de l'évaluation d'un programme similaire aux États-Unis par Alberto Abadie, Angrist, and Imbens (2002), comme l'ont également utilisé les données de Matias D. Cattaneo, Jansson, and Ma (2021) pour présenter leur nouvelle méthode. Dans cet article, les auteurs utilisent les pondérations κ de A. Abadie (2003) pour développer des effets instrumentaux quantiles de traitement et ont montré de grands effets hétérogènes selon le sexe. Pour les femmes, les effets sont généralement positifs et plus élevés au bas de la distribution des revenus, tandis que les hommes n'ont aucun effet mais à l'extrémité supérieure de la distribution des revenus. D. H. Autor, Houseman, and Kerr (2017) utilisent également ce cadre sur un programme de placement professionnel dans les emplois en premier aux États-Unis. Cependant, en se concentrant sur la distribution du revenu potentiel et non sur les effets quantiles de traitement, notre approche ne nécessite pas l'hypothèse souvent invraisemblable d'invariance du rang que de tels modèles exigent (Chernozhukov and Hansen

[32] Bozio et al. (2023) ont travaillé pour France Stratégie pour évaluer l'effet de la réforme et je faisais partie du comité de révision. Leur stratégie d'identification basée sur la différence-de-différences par nombre d'enfants a été rejetée après un test placebo sur des données de l'année précédente. Le rapport est inconclusif mais n'est pas resté dans un tiroir, ce qui est plutôt une bonne chose pour la science.

2008; Melly and Wüthrich 2017; M. Huber and Wüthrich 2019). Nos choix méthodologiques reflètent également le récent document de travail de Garbinti et al. (2023) utilisant des variations de concentration dans le temps suite à une réforme de l'impôt sur la fortune sous la présidence Hollande. À notre connaissance, cet article est le seul autre à utiliser des variables instrumentales pour estimer les masses d'accumulation et les élasticités et éviter les hypothèses paramétriques sur les densités contrefactuelles.

VII.1 Des réactions hétérogènes dans de nombreuses dimensions

Nous avons analysé les données de l'expérience pour quantifier les effets causaux du programme sur la distribution des revenus et la structure familiale. Notre principale contribution réside dans l'utilisation des variations expérimentales des probabilités d'attribution pour déduire la distribution contrefactuelle des mobilisables par variables instrumentales et mesurer les effets sur la structure familiale. En résumé, nos principales conclusions sont les suivantes :

1) **Les participantes *bunch* leurs salaires** - entre 50 % et 60 % du salaire minimum - où le taux marginal implicite d'imposition est minimal, avec peu de participants déclarant des revenus dépassant 75 % du salaire minimum à temps plein. Cette tendance est particulièrement prononcée chez celles avec deux enfants au début de l'étude.
2) Les distributions des revenus du travail parmi le groupe témoin et les non-participants ont une masse beaucoup plus faible dans la fourchette de 50 à 60 %, mais affichent une autre masse au salaire minimum à temps plein. Cependant, les disparités entre les non-participants et le groupe témoin suggèrent une plus grande probabilité de travail à temps partiel parmi les non-participants.
3) Le **taux marginal implicite d'imposition double autour de 60 % du salaire minimum**, avec des variations selon les ménages ayant un nombre différent d'enfants. Au salaire minimum à temps plein, le taux marginal implicite d'imposition dépasse 70 % pour les familles monoparentales avec un ou deux enfants.
4) Les élasticités observées pour les ménages non traités se situent entre 0,2 et 0,3, ce qui correspond de près aux estimations de la littérature existante. Cependant, l'élasticité observée dérivée des estimations paramétriques autour des points de rupture chez les participants est plus proche de 1.
5) Les estimations des densités contrefactuelles indiquent que les participantes auraient déclaré des revenus nettement plus élevés au-dessus de 75 % du salaire minimum si elles n'avaient pas participé au programme.
6) Le programme a **nettement déplacé les distributions de revenus**, reflétant des réactions substantielles à la marge intensive, et a également affecté la marge extensive pour les familles monoparentales avec un enfant au début de l'étude.
7) L'élasticité obtenue en utilisant la densité contrefactuelle des mobilisables pour estimer la masse de regroupement est d'environ 2, représentant 10 fois l'élasticité trouvée en utilisant des modèles paramétriques parmi les non-participantes.
8) Des estimations supplémentaires révèlent que la reprise d'emploi augmente moins le revenu des participantes que dans le contrefactuel, entraînant une augmentation de la pauvreté laborieuse.
9) Le programme a également augmenté la cohabitation parmi les familles monoparentales avec un enfant, réduit la fécondité parmi les familles monoparentales avec deux enfants et retardé le départ des enfants plus âgés, affectant principalement les parents de trois enfants ou plus.
10) En fin de compte, le programme n'a eu aucun effet sur le revenu disponible par unité de consommation. Les réactions et ajustements hétérogènes variés de la participation au marché du travail et de la structure familiale ont abouti à un effet nul précis sur tous les quantiles de la distribution des revenus pour tous les groupes.

Résultats principaux: réactions hétérogènes à la marge intensive La Figure 8 reproduit l'un des résultats principaux du Chapitre 3. Pour estimer les densités des "mobilisables" et des participantes traitées, nous utilisons la nouvelle méthode de distribution de régression pondérée et ajustée de l'effet à la marge extensive proposé par Matias D. Cattaneo, Jansson, and Ma (2021). Pour commencer, nous mesurons la masse de densité à un revenu de 0 pour les "mobilisables" et les participantes par variable instrumentale. Nous utilisons ensuite ces estimations pour mettre à l'échelle les densités estimées sur leur support positif. Ces estimations portent sur les revenus du travail individuels potentiels pour les "mobilisables" et les participantes. Nous utilisons des données sur une année après l'accompagnemnt et utilisons les pondérations causales pour estimer les densités contrefactuelles.

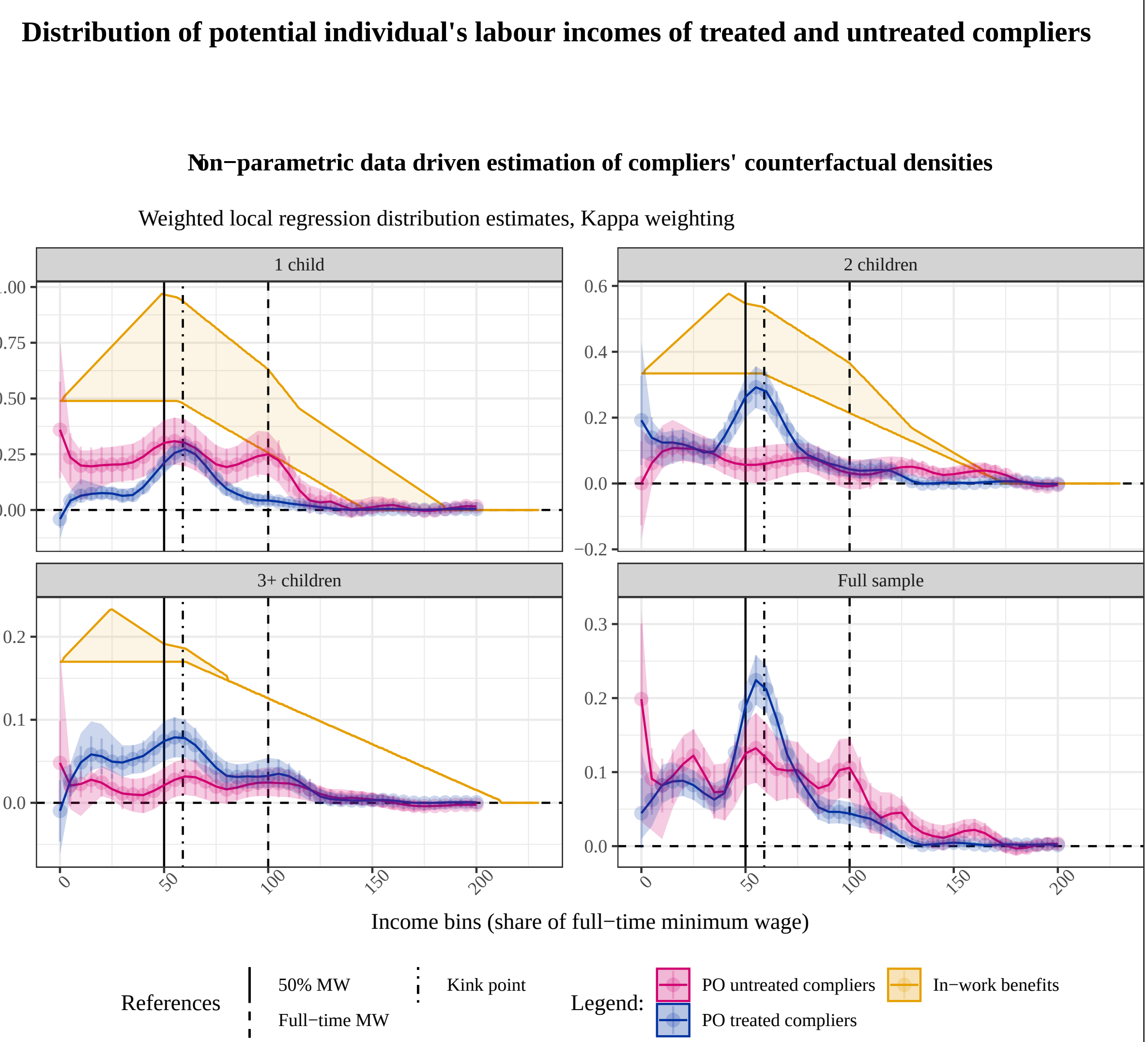
Dans les trois panneaux par nombre d'enfants, nous avons également ajouté le montant de prime d'activité empilé sur les allocations logement. Celles-ci ont été mises à l'échelle et n'ont pas d'unités verticales. Elles illustrent simplement les variations des transferts sociaux. Le quatrième panneau montre les estimations agrégées.

Les densités estimées des participantes traitées (bleu) montrent une masse de densité importante entre 50 et 60% du Smic à temps plein. À l'inverse, la distribution contrefactuelle (rose) ne présente aucune masse, pour aucun groupe. Ces densités contrefactuelles sont principalement uniformes jusqu'au salaire minimum à temps plein, à partir duquel elles chutent rapidement. L'effet le plus important concernent les mères célibataires de deux enfants au début de l'étude. Pour elles, la distribution contrefactuelle ne montre aucun signe d'optimisation. Les estimations pour les mères célibataires d'un enfant suggèrent que le programme a déplacé celles qui auraient travaillé à temps plein vers un travail à temps partiel, et a réduit la participation de celles qui auraient travaillé moins d'heures. Quant aux mères de trois enfants ou plus, nous observons des densités de revenus du travail plus élevées jusqu'au point de sortie de la prime d'activité, puis une pente plutôt linéaire.

Effets hétérogènes sur la structure des familles Enfin, notre analyse des effets du programme sur la structure familiale révèle également d'importants effets hétérogènes selon le nombre d'enfants au début de l'étude. En bref, celles avec un enfant sont plus susceptibles de se remettre en couple, celles avec deux enfants sont moins susceptibles de tomber enceintes dans l'année suivant le programme, tandis que les parents de trois enfants ou plus sont plus susceptibles de rester avec leurs enfants plus âgés plus longtemps. Ces effets sont importants et durables, montrant que les programmes actifs du marché du travail affectent de nombreuses décisions importantes allant bien au-delà de la participation au marché du travail. En fin de compte, notre résultat final confirme l'analyse du Chapitre 2 montrant un effet nul précis sur le revenu disponible par unité de consommation sur l'ensemble de la distribution des quantiles. Alors que le programme a affecté de nombreuses décisions importantes et induit de grands changements, ces réactions finissent par laisser leur revenu disponible au même niveau qu'il aurait été sans le programme, mais avec des différences critiques dans leur composition.

Cette deuxième analyse permet de mieux comprendre les résultats sur le volet extensif détaillé dans le Chapitre 2. Ce dernier montre que le programme n'a pas d'effet moyen mais inscrit les mères célibataires les plus susceptibles de trouver un emploi par elles-mêmes. Ces nouveaux résultats montrent que le programme a augmenté leur compréhension du système fiscal et d'aide sociale de manière conséquente. En conséquence, les bénéficiaires avec deux enfants travaillent à temps partiel au lieu de temps plein ; celles avec un enfant sont moins susceptibles de travailler, plus susceptibles de déclarer vivre avec leur partenaire, alignant les revenus gagnés des ménages sur la courbure de l'allocation d'activité ; celles avec trois enfants ont compris que leur niveau de transferts était principalement déterminé par le nombre d'enfants dans leur ménage, retardant le départ de leurs enfants les plus âgés et augmentant leur participation à la force de travail de sorte qu'elles ne perdent pas trop. Nos résultats indiquent également que sauf pour les mamans solos avec un enfant au début de l'étude, les bénéficiaires traitées sont plus susceptibles d'être le seul soutien de famille que les bénéficiaires non traitées, même lorsqu'elles se remettent en couple.

Figure 8: Observer l'inobservable: distributions contrefactuels et *Bunching*



Sources: ALLSTAT, observations from 18 to 30 months since random assignment.
The dependent variable is the individual's labour income as share of full−time minimum wage.
The shaded area above the estimates is the scaled amount of in−work benefits with housing benefits adjusting.
These data are simulated using EDIFIS and rescaled for each panel using the width of the density support.
Notes: Estimations of the potential density functions for compliers following Cattaneo et al. (2021), using the R package 'lpdensity'.
Bandwidth chosen using regularised integrated MSE−optimal and the Epanechnikov kernel function.
Instrument propensity scores estimated by probit of encouragement on block fixed effects to smooth support.
Densities use local polynomial regressions of order 3, simultaneous 95% confidence intervals use polynomial of order 4, 1 degree higher for biais correction.
Densities are scaled over the support of potential income with positive values. The counterfactual probability of reporting 0 income is estimated by TSLS with block fixed effects using T x 1(Y<.000001) as outcome and d instrumented by the re−centred instrument and block fixed effect, with T = D for Y(1) and T=(1−D) for Y(0).

Weights are computed for each sub−group separately: $\kappa_0 = \frac{1}{1-D_i}\frac{1-Z_i-(1-\hat{q}_b)}{\hat{q}_b(1-\hat{q}_b)}$ $\kappa_1 = \frac{1}{D_i}\frac{Z_i-\hat{q}_b}{\hat{q}_b(1-\hat{q}_b)}$

VII.2 L'*assistaxation* des familles monoparentales

Notre analyse sur le système socio-fiscal français a révélé d'importantes disparités et des désincitations à la participation au marché du travail. En particulier, les familles monoparentales bénéficiaires recevant le RSA ou la prime d'activité ne touchent aucune pension alimentaire pour leur enfant. Cette dernière est implicitement imposée à 100 % par le biais d'une baisse des prestations sociales, réduisant du même montant toute incitation financière à l'emploi à moins de gagner suffisamment pour ne plus être éligible à la prime d'activité. Alors seulement, percevraient-elles la pension alimentaire qui leur était pourtant versé. Toutes les autres prestations familiales sont également entièrement implicitement imposées pour les bénéficiaires de minima-sociaux, tandis qu'elles ne font pas partie du revenu imposable pour les parents qui n'y ont pas recours. De plus, nos simulations révèlent que le taux marginal implicite d'imposition pour les familles monoparentales dépasse celui des couples dans tous les scénarios que nous avons explorés, avec une fourchette entre 50% et 60% du SMIC à temps plein représentant le taux d'imposition implicite le plus bas, quel que soit la structure du ménage. Cependant, celui-ci est borné par un taux marginal d'imposition implicite environ deux fois plus élevé. Au-delà du salaire minimum à temps plein, les familles monoparentales avec un ou deux enfants font face à un taux marginal d'imposition implicite dépassant 70%. Sans le savoir, celles qui demandent le RSA signent en même temps pour le plus haut taux de taxation de toute la distribution de revenu.

les familles monoparentales, en particulier avec plusieurs enfants, font face à un taux d'imposition implicit particulièrement élevé, notamment aux seuils de revenus juste au dessus du SMIC à temps plein. Cette situation crée des désincitations fortes à augmenter la participation au marché du travail et entrave la sortie de la pauvreté par l'emploi. Les interactions complexes entre diverses prestations sociales contribuent à un manque de transparence et de compréhension parmi les bénéficiaires, entravant leur capacité à prendre des décisions éclairées.

Nous avons proposé le terme "*Assistaxation*" pour désigner ce phénomène consistant à taxer massivement les ressources économiques, physiques et mentales des personnes ayant recours à l'aide publique, leur laissant au passage peu de moyen de s'en extraire. Il ne s'agit pas seulement du fardeau administratif et de la stigmatisation discutés dans la littérature ; l'assistaxation implique également une imposition plus lourde et implicite des revenus du travail et un taux d'imposition de 100 % sur la pension alimentaire et des prestations familiales des plus démunis. L'*assistaxation* reflète également mieux le type de taxation qui est évité ici : un régime sophistiqué et des mesures coercitives pour forcer la participation au marché du travail tout en les taxant le plus, à moins qu'ils ne quittent l'aide ou ne se mettent en couple.

Faire des parallèles avec le contexte historique des Poor Laws révèle des similitudes frappantes dans les inégalités systémiques et les obstacles à l'avancement économique auxquels sont confrontés les groupes marginalisés (Persky 1997). Tout comme les Poor Laws ont enraciné la pauvreté et perpétué la stratification sociale, l'assistaxation perpétue les cycles de difficultés financières et d'inégalités, entravant au passage tout effort déployé pour s'extraire de cette situation.

Notre travail met en évidence d'importantes contradictions dans le système socio-fiscal, ainsi que dans les discours politiques. Pour les familles monoparentales, les transferts sociaux sont à la fois trop faibles pour réduire la pauvreté et s'épuisent trop rapidement pour *"rendre le travail payant"*. Pourtant, la plupart d'entre elles l'ignorent. Le gouvernement a très peu d'incitations à rendre les règles plus claires, car les réactions induiraient très probablement une participation et des heures de travail plus faibles des mamans-solos, augmentant ainsi davantage la pauvreté laborieuse, les inégalités de genre sur le marché du travail et la pauvreté chez les enfants. De l'autre côté de la distribution des revenus, Garbinti et al. (2023) ont analysé l'effet d'une réforme de l'impôt sur la fortune en France et ont également constaté de fortes réactions - une moindre richesse pour les contribuables mieux informés. Cependant, ces résultats montrent qu'ils proviennent principalement d'évasion fiscale réelle et de déclarations manipulées facilitées par le nouveau régime fiscal. En revanche, les familles à faible revenu font l'objet d'un contrôle plus étroit grâce au ciblage algorithmique. Les contrôleurs ont accès aux comptes bancaires des bénéficiaires et désormais, l'administration reçoit des flux directs des taxes sur les salaires, rendant les manipulations de revenus presque impossibles et très dissuasives. Cela entraîne également une grande perte de liberté et de droit à la vie privée, ainsi que des taux d'imposition implicites plus élevés (Quadrature du net 2023; Défenseur des droits 2017). Face à des conditions d'emploi plus difficiles et à des emplois de mauvaise qualité (Rodrik and Stantcheva 2021), les familles monoparentales bénéficiaires de l'aide sociale semblent vraiment piégés dans une spirale de pauvreté dont il est difficile de s'échapper sans changement radical de politiques publiques.

En fin de compte, il y a une question fondamentale sur les préférences en matière de redistribution, de justice sociale et de taxation optimale (Blundell et al. 2009; Maniquet and Neumann 2021; Saez and Stantcheva 2016; Stantcheva 2021). Des travaux récents montrent que les préférences sociales sont fortement polarisées politiquement mais changent avec une meilleure connaissance (Kuziemko et al. 2015; Alesina, Stantcheva, and Teso 2018; Hvidberg, Kreiner, and Stantcheva 2023). Un moteur clé des préférences sociales tourne autour des inégalités d'opportunités et notre travail démontre que le système socio-fiscal crée et exagère certaines d'entre elles. Nous espérons qu'en dénonçant ces injustices, nous pourrons favoriser des réformes qui permettront enfin aux familles monoparentales de gagner en autonomie, moyens et liberté pour améliorer leur vie et celle de leurs enfants.

Part 1

Social investments in early childhood: The daycare assignment problem

Chapter 1.

Rage against the matching: fairness and inequalities in a market design experiment of daycare assignments in France

This Chapter is co-written with Julien Combe. It is part of our long-term research project called ISAJE *Investissement social dans l'accueil du jeune enfant* joining efforts between French National Family Allowance Fund (Cnaf) École polytechnique and Paris School of Economics. This work would not have been possible without their financial, human and administrative support. We specifically thank Saad Loutfi for his great work on the various assignment waves and Jeanne Moeneclaey, Virginie Gimbert, Florence Thibault, Bernard Tapie, Lucie Gonzalez and Vincent Mazauric for their trust and support throughout the project. We also thank Olivier Noblecourt and Jean-Benoit Dujol for their political support. We owe a great deal to the *Conseil d'agglomération* of Valence Romans Agglo but above all to its Early Childhood Department, which allowed us to start this experiment, enabled this data analysis, promoted our work to other areas and gave us its unfailing confidence. Special thanks go to Julie Vivant and her team, with whom we had an excellent working experience. We also thank the early childhood department and incumbents of the other cities that participated in this ambitious experiment. We are grateful to Anne Boring, Karen Macours and Camille Terrier for helpful comments on a first version. Special thanks to Marc Gurgand for is constant and empowering support. We would like to thank Pierre Boyer, Antoine Bozio, Julien Grenet and Delphine Roy and the Institut des Politiques Publiques (IPP) - Paris School of Economics for providing free-of-charge additional staff during this experiment. Special credits and gratitude go to to Paul-Emmanuel Chouc and Vincent Verger (IPP), for their fantastic work on various assignment waves, Agathe Eupherte, Raja Ahmed Taleb and Marion Goglio for their remarkable research assistant work. We also thank the members of the ISAJE scientific committee for helpful guidance through this work and especially Orla Doyle and Sylvana Cote. This work greatly benefited from the careful proofreading and valuable feedback provided by Alex Galitzine, to whom we express our sincere gratitude. Final thanks go to Laudine Carbuccia, Quentin Daviot, Marine De Montaignac, Cécile Ensellem, Gautier Maigne, and many other colleagues and friends for helpful discussions and support over the years.

Abstract

The choice of childcare alternatives is a central decision which affects several key societal dimensions such as child development, mothers' labour supply, and economic and gender inequalities. While families from lower socioeconomic backgrounds tend to gain the most from formal childcare, these services are overwhelmingly used by more affluent families - a phenomenon often called the *"Matthew effect"*. In this paper, we consider access to daycare as a matching problem controlled by local authorities. Based on policymakers' definitions of the procedure, we use market design to define assignment mechanisms and we analyse the consequences of important design choices in a field experiment. The daycare assignment problem is similar to school choice, but includes two additional features: multidimensional constraints that cover weekdays and diversity constraints, typically age groups. Policymakers' design choices affect the definition and range of stable matchings. Our algorithms deliver student optimal *fair* assignments (SOFA) in the different versions of the problem. From 2020 to 2023, we assigned daycare slots to families in nine urban districts in France. Our objectives were twofold: i) to provide automated assignment mechanisms with desirable properties and ii) to introduce random variation in assignment probabilities for future work on causal effects of accessing daycare. We use two case studies to demonstrate that our assignments meet their intended objectives and compare different assignments. Using a change in priorities in one case study and counterfactual simulations of alternative priority scores, we show that i) giving larger weights to some group (e.g. dual earners) increases their assignment probabilities and share in daycares, ii) increasing priorities with time since registration strongly penalises single parents, in part because iii) dual-earner couples compete for early entry and strategically register as soon as possible, as incentives. iv) Being the largest group and also receiving high social weights, their strategies crowd-out parents who cannot register that early and create inequalities of opportunities even among strategic parents, correlating assignment probabilities with birth month. Other analyses show that diversity constraints may create sharp discontinuities in assignment probabilities that are unrelated to priority scores. Our results provide clear evidence of the mechanisms that contribute to the Matthew effects in childcare, and they are mostly political choices. However, our tools can be used to satisfy other distributional objectives and achieve higher transparency in the assignment process.

- **JEL Classification Numbers** : D47, A13, D82, D63, D78, A13, J13, I38, H42
- **Keywords**: Market design, fairness, childcare, daycare assignment, early childhood, inequalities, social investment, France

Résumé

L'accueil des jeunes enfants affecte plusieurs dimensions sociétales clés, telles que les inégalités de développement de l'enfant, la participation au marché du travail des mères, ainsi que les inégalités sociales, économiques et de genre. Alors que la recherche souligne que ce sont les familles issues de milieux socio-économiques moins favorisés qui bénéficient le plus des modes d'accueil formels, ce sont surtout les plus favorisées qui y ont recours - un phénomène souvent appelé *"effet Matthieu"*. Dans cet article, nous considérons l'accès en crèche comme un appariement centralisé organisé par les autorités locales. En nous basant sur les définitions du problème par les décideurs politiques, nous proposons des modèles pour définir des mécanismes d'affectation et leurs propriétés. Nous analysons ensuite les conséquences des choix de conception dans le cadre d'une expérience de terrain. Le problème est similaire au choix d'école, mais comprend des contraintes multidimensionnelles - comme les jours de semaine -, et des contraintes de diversité - typiquement des groupes d'âge. Les choix des décideurs politiques affectent la définition et la gamme des appariements stables. Nos algorithmes fournissent des affectations optimales équitables pour les familles (SOFA) selon la définition des décideurs. De 2020 à 2023, nous avons affecté les places en crèche de neuf larges collectivités territoriales en France. Nos objectifs étaient doubles : i) fournir des mécanismes d'affectation automatisés aux propriétés désirables et ii) introduire une variation aléatoire dans les probabilités d'affectation pour des travaux futurs sur les effets causaux de l'accès en crèche. À partir de 2 études de cas, nous montrons que nos affectations ont les propriétés attendues, et comparons les alternatives et ajustements en commission. Nous nous appuyons sur une réforme des priorités dans une étude de cas pour simuler des affectations contrefactuelles avec des scores de priorité alternatifs. Nous montrons que i) donner plus de poids à certains groupes (couples bi-actifs en particulier) augmente leurs probabilités d'affectation et leur part dans les crèches, ii) inclure l'ancienneté de la demande dans les priorités pénalise fortement les familles monoparentales, en partie parce que iii) les couples bi-actifs sont en compétition pour les entrées précoces et s'inscrivent stratégiquement dès que possible, conformément aux incitations. iv) Étant le groupe majoritaire et parmi les plus priorisés au départ, leurs stratégies évincent les parents qui ne peuvent pas s'inscrire aussi tôt et créent des inégalités d'opportunités même parmi les parents stratégiques, corrélant les probabilités d'affectation avec le mois de naissance. D'autres analyses montrent que les contraintes de diversité peuvent créer des discontinuités nettes dans les probabilités d'affectation qui ne sont pas liées aux scores de priorité. Nos résultats documentent clairement les mécanismes qui contribuent à un effet Matthieu dans l'accueil du jeune enfant, et ils sont principalement des choix politiques. Cependant, nos outils peuvent être utilisés pour satisfaire d'autres objectifs de distribution et atteindre une plus grande transparence dans le processus d'affectation.

- **Codes Journal of economic literature** : D47, D82, D63, D78, A13, J13, I38, H42
- **Mots clés**: Design de marché, petite enfance, crèche, inégalité, investissement social, France.

I Introduction

Public provision of daycare is a central tool for policymakers, particularly within the European Union. Early childhood care and education policies are framed as crucial social investments[1] aimed at breaking the cycle of poverty, addressing gender inequalities, enhancing social mobility, and boosting labour market productivity (Morel, Palier, and Palme 2012a; Van Lancker 2013; Kvist 2015; A. Hemerijck and Huguenot-Noël 2022). These goals rely primarily on the externalities of early childcare policies through both the technology of skills formation and parents labour market outcomes (Connelly and Kimmel 2003; Tekin 2007; Bauernschuster and Schlotter 2015; Zoch 2020; Doyle 2020). The positive impact of access to childcare services, particularly on children from disadvantaged backgrounds, has been documented in numerous studies[2]. However, the use of these services is skewed in favour of higher-income families (Lancker and Ghysels 2012; Petitclerc et al. 2017). This phenomenon is known in the social investment literature as the "*Matthew effect*" and raises important questions whose answers have different policy implications. Is the *Matthew effect* due to factors on the demand side ? Is-it because of local demographics ? Parents preferences ? Do they depend on cultural norms around parenthood ? Or is is because of supply-side constraints such as limited provision, costs or institutional settings ? So far, evidence suggests that supply side constraints play a larger role in explaining the Matthew effect than demand side factors (Farfan-Portet, Lorant, and Petrella 2011; Abrassart and Bonoli 2015; Pavolini and Van Lancker 2018 ; Carbuccia, Thouzeau, et al. 2023).

This study is about what lies in between: *how* parents access formal childcare, and what policymakers can do to effectively reach their ambitious goals. While there is a large and growing body of research on the benefits of early childhood interventions, the *engineering* of childcare policies does not always yield the high and long term returns science promised (Cantillon 2011; Ghysels and Van Lancker 2011; Vandenbroucke and Vleminckx 2011a). Our main hypothesis is that the *Matthew effect* arises, partly at least, because markets for childcare are rationed and not properly designed to account for the externalities of childcare. Because of rationing and heterogeneity in the effects of childcare on different populations, "*who gets what and why*" matters (A. E. Roth 2015).

In this paper, we consider access to daycare as a matching problem where demand meets supply through a centralised assignment mechanism organised by local authorities. Based on policymakers' definition of the problem, our main research questions are:

To what extend is an optimal daycare assignment mechanism possible ? What are the consequences of different design choices ?

Our main contributions are twofold. First, we provide models and algorithms for different versions of the daycare assignment problem with well understood properties. Applications provide policymakers new tools to assign parents their most preferred daycare centre following priorities with diversity and multidimensional constraints. Second, we use two case studies to analyse the effects of diversity constraints, priorities and assignment mechanisms on inequalities in daycare access across different dimensions.

At the onset, we developed these models and collected these data to answer completely different research questions. Indeed, we wanted to run a randomised experiment to measure the effect of accessing daycare on children development, mothers' labour market participation and so on. However, one does not simply fully randomise access to formal childcare when assignment is centralised by local authorities[3]. In addition to the usual barriers to experiments, there are many different providers over which parents have preferences, local authorities are in charge of the assignment and have policy objectives of their own. There is simply no standard experimental design that easily accommodates these constraints. Instead, we followed Abdulkadiroglu et al. (2017) and made our *"research design meet market design"*: we modelled and automated daycare assignment procedures with an embedded randomised

[1] The European agenda embraced the term in 2013 with the introduction of the "Social Investment Package" (SIP). This framework aimed to redirect Member States' social policies toward lifelong social investment. The SIP defined social investments as measures that "strengthen people's skills and capacities and enable them to participate fully in the labour market and society. Priority areas include education, quality childcare, healthcare, training, employment support, and reintegration" (Commission 2015). This definition underscores the core principles of social investment: investing in individuals' human capital and promoting labour market participation (Deeming and Smyth 2015)

[2] See for instance Nores and Barnett (2010), Barnett (2011) or Kholoptseva (2016). More recently van Huizen and Plantenga (2018) systematically synthesised quasi-experimental literature assessing the effects of access to childcare services between 2005 and 2017 across developed countries.

[3] Although it has been done in Rio (Brazil) and the effects of the reform have been studied by O. Attanasio (2022).

experiment in the *plumbing*. By doing so, we minimised disruption to the ecological conditions of access to childcare, we garnered support from territories without encroaching on local policies, putting us in a unique position of *marketmaker* and *scientist* to better understand market structures, forces and consequences. Last but not least, demands with the same priority levels were randomly ordered in the assignment procedure, generating experimental variations for future causal analysis.

In 2019, we started a long-term research programme and partnership called ISAJE (Social Investment in Early Children's Lives and Education) with the aim of providing high-quality evidence on the effects of childcare policies. From 2020 to 2023, we convinced nine large urban municipalities and local authorities to participate in an experiment that would automate their daycare assignment procedure with random tie-breaks[4]. One territory stayed for four years and provided data for most of the empirical part of this research. Two others participated for three years, three stayed for two years, and three others for one year. Ultimately, approximately 20 000 demands went through our automated procedures for 5 000 daycare slots. Within these 20 000 demands, between 2 000 and 3 000 applicants were subjected to a lottery. The next steps of this research project will be to match these databases with monthly administrative records from the National family allowance fund (Cnaf) on all parents of children under 3 in the area and build an administrative panel for the next ten years.

In this research, it is important to note that our main goal was to be able to use the data for subsequent causal evaluations. Therefore, our theoretical contribution is incremental and mostly based on existing, well defined procedures. In practice, we met policy-makers, daycare providers, heads of early childhood departments in local councils and so on. We asked what they were trying to achieve with their assignment procedure and how they were proceeding. We noticed that the French daycare market had some familiar features. In particular, i) assignments are centralised by local authorities, ii) parents register throughout the year and submit preferences over daycare centres, iii) most offers are available in September when older children move to preschools and iv) local governments organise a main assignment committee during spring to assign these slots.

The daycare assignment problem is therefore very much like static school choice problems for which well defined solutions and many successful applications already exist (Abdulkadiroğlu and Sönmez 2003). In this literature, the goal is typically to define *stable* matchings and algorithms to find them. *Stability* is a property that we want a matching to respect in order to justify to parents *why* the assignment is as it is. The theoretical motivation for concentrating on sets of stable matchings is that if the market outcome is unstable, there are agents or pair of agents who have the incentive to circumvent the match (A. E. Roth 2002). In some instances, scholars and/or policymakers already use results from the market design literature to provide mechanisms for childcare assignment (see our review in section II).

However, the definition of stability highly depends on the constraints of the problem. As Kominers, Teytelboym, and Crawford (2017) note, "*successful market design solutions are bound to vary across markets because real-world settings have distinct (and sometimes unexpected) objectives, constraints, and trade-offs*". For a given market and definition of stability, a stable matching may not exist, or there can be several. There is a range of possible designs that can produce very different *outcomes*, *i.e.* in terms of the distribution of assignments, as well as a set of *consequences*. As Li (2017a) writes, a consequence is richer than an outcome, it is a "*description of the effects of the market on the world*". Designing marketplaces for daycare requires weighing and solving the trade-offs between different objectives and constraints, and models and theory help to make clear what (some of) those trade-offs are.

In our settings, many local institutions collect preferences over week days and assign files accordingly[5]. This simple change in the way preferences are collected have important consequences on the definition of stability. Policymakers are familiar with the details of their environment, and yet often do not know how to state their objectives in precise terms or realise they are conflicting. However, market design provides "*guidance without prescribing an entire ethical theory*" (Li 2017a). Many policy objectives depend on what is feasible. For instance, a policymaker might think that it is morally obligatory to implement a system that is Pareto optimal[6] for families and respect pri-

[4] Importantly, local authorities were contacted for this purpose. The research was about evaluating the impacts of childcare, not market design. Most of those who participated showed strong interest in these research questions and understood what causal evaluation required and accepted randomisation as a condition for participation.

[5] In the absence of specific days, one child takes one seat, which is an indivisible and unitary good. However, in settings with specific days, a student can take some days and the rest may be assigned to another student with complementary preferences, or not. A seat is no longer indivisible and unitary.

[6] In this context, a Pareto efficient assignment implies that the distribution of childcare slots cannot be improved in a way that benefits at least one participant without harming another participant.

orities. But this is not possible (Abdulkadiroğlu and Sönmez 2003) so it cannot be a moral or political constraint. When assignment is done by allocating slots over weekays, recent developments on matching with multidimensional constraints for childcare (Kamada and Kojima 2023) and refugee resettlements (Delacrétaz, Kominers, and Teytelboym 2023) provide different definitions of stability and algorithms to find stable assignments with other desirable properties. Importantly, these papers show that no algorithm can respect priorities[7] and be non-wasteful. Again, both cannot be achieved so it cannot be a policy objective. This recent literature provides well fitted theoretical solutions for the daycare assignment problem with multidimensional constraints.

However, real life implementation adds another layer of complexity. Indeed, another important feature in the French system is the existence of diversity constraints in all childcare centres and assignment procedures. Diversity constraints imply that capacities within daycare centres are divided into *buckets*, *i.e.* bundles of capacities with attached priority rules over *groups*. For instance, some capacities will be reserved for children aged 6-12 months only, and others for children aged 12 to 24 months. These diversity constraints can be *hard* when buckets only accept one group, like in the previous examples. Sometimes, constraints are *soft* and define a preference order over groups. For instance, a bucket may accept children aged 12 to 24 months but also older children if there is enough capacity to accommodate them. While the most common forms of diversity constraints are age groups, policymakers and/or daycare providers define many different forms of diversity constraints[8].

These diversity constraints play an important role in the lack of transparency in the daycare assignment procedures. Indeed, assignment committees do not simply sort files by priorities in a daycare, they consider assignment within buckets, which definitions vary within and across daycare centres, generating more or less competition across groups and thus, variations in assignment probabilities. In a given matching, there may be children with low priority scores assigned to a daycare that higher priority parents wanted but were rejected from. Thus, the stability notion cannot be solely based on priority scores, it must account for the partition of capacities into buckets and priorities within and across buckets.

Providing a stability notion in a case with multidimensional and diversity constraints is therefore both relevant theoretically and necessary to ensure transparency and due process.

Our theoretical contributions We define daycare assignment marketplaces (DAM) as three elements policymakers must choose :

1) A version of the problem : whether they consider demands with specific days or not ;
2) A partition of capacities in *buckets* i.e. bundle of capacities with attached priority rules in each daycare to define diversity constraints ;
3) A definition of *fairness* : with priority scores and whether the mechanisms should eliminate justified envy, tolerate some small deviations or consider initially feasible demands only.

Building on the work of Ehlers et al. (2014) in a case without day slots assignment, Kamada and Kojima (2023) and Delacrétaz, Kominers, and Teytelboym (2023) in a case with day slots assignment, we propose stability notions for different versions of the daycare assignment problem with diversity constraints. Our main result – presented in Theorem 2 – then states that for each DAM, we can find the unique student optimal fair assignment (SOFA), that respects the *chosen* definition of envy-freeness. We also provide an additional definition of stability over *initially feasible demands*. From a SOFA with all demands, we show in Theorem 1 that removing initially infeasible demands is weakly preferred by all parents and thus, Pareto-improves the assignment they receive. In practice, this theorem proves most useful in second rounds where there are many more empty days in buckets.

7 with different definitions in the two papers.

8 e.g. for parents whose schedules vary over the months or with late-night shifts, those in an active labour market policy involving childcare arrangement (*crèche AVIP*), quotas for children with disabilities, etc.

Field experiment The second part of our research focuses on the practical application of these automated daycare allocation mechanisms. In this study, we did not have a direct hand in defining the priority criteria, and in some instances, we were not privy to the exact methodology used to calculate scores. At the time of writing, confidentiality agreements prevent us from using all data. However, we have anonymised data sharing agreements with two local administrations. We primarily use data from one large urban centre for which we have four years of operation. Importantly, the datasets are rich enough for a comprehensive understanding of the criteria, weights, and formula for priority score, and encompass a reform of the latter in the middle of the time frame. This mechanism is based on weekday demands with flexible constraints over age groups. The second case study is only one assignment based on the school choice version of the problem. The datasets contain less informations but this setting is interesting for its large discrepancy in relative supply across age groups.

Our findings are based on four key empirical assessments strongly inspired by the 4-step auditing of algorithm proposed by Kasy and Abebe (2021). First, we examine the properties of our algorithms and compare the outcomes of various mechanisms with the final assignment. Second, we leverage a reform of priorities in the middle of the time frame of Case Study I to measure its effects and that of alternative priority rules on inequalities and segregation between social groups. Third, we investigate the impact of strategic registration and its consequences on inequalities by birth month. Fourth, we use data from Case Study II to demonstrate the critical role of bucket elements and distribution.

Using data from the 2023 rounds, we first start by showing that our assignments are indeed family optimal and envy-free and usually not *too wasteful*. Then, we compare outputs of different algorithms with the final offers sent to families. This year, the committee followed our proposed assignment in 92.3 % of all cases, so there are deviations. A significant share comes from the inability to ensure joint assignments of children from the same family in the same facility. This is one of the main drawbacks of this assignment procedure. However, these slight deviations go beyond grouping children of the same family: they unmatch and move few files to also increase the number of children assigned. We find this paradox particularly interesting because while we did offer to use a weaker definition of envy-freeness to increase the number of seated children, they refused and opted instead for hand-made deviations to achieve a similar goal.

In this market, policymakers took great care in defining a multi-criteria priority score accounting for children's and parent's social situations and needs, family structures and employment, in particular. But parents do not know precisely how it is built, nor the value of their own scores. The public documentation enumerates the main criteria including time since application and encourages parents to contact the early childhood department as soon as the pregnancy is official[9].
The importance of time since registration in priorities is such that any disturbance in the distribution of registration or timing of birth can have important effects on the distribution of assignments within and across social groups. In this experiment, the Covid-19 pandemic caused delays in both registrations and birth, reducing variations in priorities as well as the share of satisfied double earner couples. In reaction, policymakers boosted priorities of parents in employment such that in 2022, only 30% of single parents with no job had higher priority scores than the 10% dual earners with lowest priorities.

We use our tools to simulate assignments with the former weights, no priority for time since registration and no priority at all and compare average assignment change by social group from individual pair difference between actual and counterfactual assignment. These *what if* scenarii show that the reform worked: it increased the proportion of dual-earner couples and active single parents at the expense of other demographic groups and increased segregation within daycares. It also shows that without time since registration, conditional assignment probabilities across group would have been roughly the same, except single parents who are disproportionately hurt by this criteria. In this setting, the other weights mostly reduce assignment probabilities of mothers without a job. For instance, if there were no priorities in the 2023 round but only preferences, capacity constraints by age and random ordering of applicants, the assignment probabilities of single-earner couples and single parent with no job would be 12 pp higher, while dual earner couple's would be 7pp lower. Such simulations offer valuable insights for policymakers, facilitating the development of more equitable weighting schemes.

Because of labour constraints, dual earner couples often require and decide their childcare arrangement earlier. In this setting, they represent 66% of the demand and all receive the second highest social weights. Other criteria affect

[9] The declaration of pregnancy to the social security system is generally made by the doctor (general practitioner or gynaecologist) or midwife before the end of the 3rd month of pregnancy. Note that not all assignment committees allow parents to register that early.

few individuals and time since registration explains more than 3/4 of variations in priority. Early registration is a dominant strategy bounded by declaration of pregnancies. Consistent with these incentives, we observe massive bunching of registration 5 months before birth for dual earner couples. However, such a strategy is out of reach for parents whose situation changes later (separation, lay-offs, hiring and so on) or simply for those who think of childcare later during their pregnancy. Consistently, there is no early bunching among single parents and other couples. While these two features already strongly favour stable couples in employment who anticipate their needs, policymakers give dual earner couples the second highest social weight, and down-weight families where at least one does not work, further accentuating social inequalities in daycare access.

Moreover, waiting for the decision of the assignment committee requires alternative childcare arrangement until they are accepted, and outside options if they are not, especially for active parents. Paradoxically, the more they wait, the more likely they are to be seated and the more likely they are to refuse the committee's proposition. Ultimately, congestion is artificially inflated by parents incentivised to register early and anticipate their needs months before birth and very far away from accessing daycares. By the time they receive an offer, many have changed their mind, or wait for a complete choice set to make a decision. These parents impose negative externalities on rejected less prioritised parents and cause part of the market to unravel.

Because pregnancies occur throughout the year but assignment committees occur in April, early-registration does not have the same pay-off for strategic parents who get pregnant in spring or autumn. The conditional assignment probability by registration month depends on the density of applications, which is lowest in the summer months, mostly for children expected in the last quarter. Therefore, conditional on registration, children born in December have the highest assignment probability of all.

However, other rules may also create variations in assignment probabilities by birth months which may favour children in other seasons. In particular, the definition of the bucket plays a pivotal strategic role in shaping assignment distributions. In Case Study II, most seats are only open for children born after the 1st January. Using a regression discontinuity design (Calonico, Cattaneo, and Titiunik 2014), we show that this sharp difference in available supply across age groups creates a sharp discontinuity in the assignment probabilities, but not in priority scores or density of the forcing variable. Our findings indicate that the probability of assignment for registered parents is heavily influenced by the synchronisation of birth dates, but through distinct mechanisms, here precisely reversed. In this setting, time since registration did not affect priorities, but policymakers supplied fewer seats for children born in the previous calendar year, As a result, children born in the first two quarters had higher assignment probabilities, while those born at the end of the year experienced much lower probabilities. In the first case study, unequal access was mostly caused by strategic early registration by the largest group, which increased their priority and resulted in higher assignment probabilities for children born in the last quarter who were registered around three months after conception. Conversely, children born in the first two quarters experienced lower assignment probabilities.

Policy implications Thinking of the market structure for childcare helps understanding *why* the assessment of universal childcare policies doesn't always yield the anticipated outcomes and *how* the socio-economic gap in early childcare enrolment occurs. In our examples, the Matthew effect is clear and stems from political choices on key design elements. We discuss this aspect more thoroughly in Section VI.

Once policymakers define the elements of the DAM, we can provide parents with their favourite choice among assignments, all the while respecting priorities and diversity constraints following the definition of *fair* assignment. However, this definition of *fairness* is very narrow (Kasy and Abebe 2021). It only encompasses *procedural justice*, *i.e.* due process and equal treatment of individuals. Yet the perception of justice encompasses other aspects and in particular, *distributional justice*, *i.e.* the "*allocation of positive and negative outcomes in a decision context and whether they are distributed equitably or deservedly amongst the affected population given their circumstances, performance or contributions*" (Binns et al. 2018). Perceived levels of justice of a decision outcome are separate from purely self-serving rationalisations of a decision outcome; *"an individual might be negatively affected by a decision whilst still thinking it is just"* (Binns et al. 2018). Conversely, if priority rules are not perceived as fair, then neither will the assignments that respect them (Fenech and Skattebol 2019).

The initial experimental phase provided an opportunity to demonstrate the relevance and effectiveness of automated procedures for different versions of the daycare assignment problem. They can also be used to offer new possibilities. For instance, it is easy with buckets to allow parents in a daycare to try and change for another one without

derogatory rules These solutions have proven effective in automating allocation committees while ensuring a realistic but controlled environment for a "natural" randomised experiment. However, the willingness to participate in the research does not reflect the general acceptability of market design solutions. In this experiment, the automation often occurred with minimal information provided to families, primarily for reasons related to the scientific integrity of the research protocol[10]. However, the *official* adoption of our DAM is likely to elicit different reactions from families and childcare providers. This point relates to an important aspect of market design: *repugnance* (A. E. Roth 2007; Satz 2010b).
Disgust for certain types of transactions can pose a genuine constraint on markets and their design, similar to technological constraints or requirements for incentives and efficiency. In the early childhood care market, there are transactions that are considered unethical, such as attempting to win favours with elected officials in order to be selected. Ensuring envy-free assignments is a way to reduce such repugnant transactions[11].

In certain contexts, some market structures may be more objectionable than others. The very idea that access to childcare can be determined by something other than a human choice can be particularly startling to some people. In general, the combined use of priorities and lotteries in an automated process may be seen as a departure from traditional decision-making mechanisms, potentially challenging established trust, power dynamics and political considerations. Parents of different social background have different views on fairness and social justice (Hvidberg, Kreiner, and Stantcheva 2022).

By ensuring procedural justice, our daycare assignment mechanism moves fairness concerns over *distributional justice*. Ensuring transparency may only be feasible if the definition of the elements of the DAM align with societal values and reflect social justice principles[12]. Importantly, the definition of buckets is part of the definition of envy-freeness and relative supply constraints may be more important than priority scores in their distributional effects. When we initiated the project, we faced strong resistance from the French Mayors' Association (AMF). They saw this as a potential interference in their rights and power as elected representatives. Others objected to the project because they feared that automating processes would put jobs in local early childhood services at risk or that daycare providers would lose their bargaining power.

The adoption of such technology is therefore not merely an issue of optimisation, efficiency, or even transparency: it implies considerations of equal opportunities, social justice, and the functioning of democracy. In this regard, scaling up raises the question of societal reception to the adoption of a new technology and the conditions under which acceptance, mistrust, or rejection may occur. Before thinking of scaling up, perhaps the first question a democracy should ask is: "*who gets to pick the objective function of an algorithm?*" which is intimately connected with the political economy question of who has ownership and control rights over data and algorithms (Kasy and Abebe 2021; Schmauder et al. 2023; Albright 2023; Bohren, Hull, and Imas 2022; Kasy 2023).

The rest of the paper is structured as followed. In section II, we briefly describe the French market for childcare and review the international literature on daycare assignment. Section III presents the experimental framework and main features of our two case studies. Section IV presents our models for daycare assignment mechanisms. We start with a simple school choice model, then introduce multidimensional constraints with solutions from Kamada and Kojima (2023) and Delacrétaz, Kominers, and Teytelboym (2023) and finally we define daycare assignment mechanisms with diversity constraints. Section V is our empirical analysis of a daycare assignment mechanism over four years. We conclude section VI on the implications of our work and future perspectives.

[10] Parents were informed that the city hall took part in a research on early childcare as bare minimum.
[11] We had informal feedbacks from some civil servants of several places that direct interventions from elected officials were dramatically reduced since the adoption of our assignment procedure.
[12] For a debate in feminist economics on these issues, see *e.g.* Albelda, Himmelweit, and Humphries (2004) or Thomson (2009).

II Families, policies and markets for childcare

II.1 The market for childcare in France

The French system offers parents a range of support measures, including parental leave, tax credits for childcare spending, family benefits, publicly or privately funded daycare centres, and subsidies for childminders. In 2020, the French government spent € 6.6 billion for collective childcare, € 4.6 Billion for individual care (childminders), € 1.7 billion for tax credit (Ishii et al. 2023). If all of these resources were fully utilised, centre-based childcare could take 471,000 children (20% of children under three), and childminders 744,000 (33% of children under three). Overall, the total coverage rate of childcare for children under three is about 60%, and the total direct public spending on childcare in 2020 was €14.3 billion. In Appendix A, we describe the main types of available childcare services along with the aids for families with children under 3.

Early-childhood policies Formal childcare services are accessible to children as early as three months of age, and preschool education has been made mandatory by the age of three years in 2019[13]. Table A.2 presents the evolution of the main childcare arrangements for children under 3 from 2002 to 2021 based on representative survey reported in Caenen and Virot (2023). 20 years ago, 70% of children were mainly cared for by parents but now, almost half of all children under three are in other childcare arrangements. Childcare centres saw their share grow, now serving 1/5 of children under 3. There is a strong rationing in access to centre based childcare, 40% report daycare as their favourite choice but only 25% access one (Laporte, Crépin, and Hilairet 2019). Childminders face fewer constraints, with 1/3 of families considering them their favourite and the same proportion using their services. Moreover, 40% of families combine various solutions, while 56% rely on parental care, essentially mothers out of the labour force.

Local authorities manage access to public daycares The process of allocating daycare slots is primarily overseen by local authorities, with limited regulation or guidance on the assignment process. The law only mandates higher priority for applicants with disabled children, those in social and professional integration processes, and families referred by social services[14]. For the majority of applicants, each municipality determines their own criteria. Decisions are typically made by committees comprising elected officials, daycare managers and civil servants from early-childhood services.

Allocation committees function as centralised institutions that bring together applicants and daycare providers, operating as marketplaces designed to match supply and demand. While prices, or childcare costs, are relevant factors in these marketplaces, they do not solely determine daycare placements. Decision-making involves not only applicants but also local agencies and childcare providers, who have their own preferences regarding which children to accept. Parents also have alternative options, such as private childcare or taking care of their children themselves. To our knowledge, Herman (2017) is the only researcher, besides ourselves, who analysed the daycare assignment procedures in France. She reports three monographs and compare the different organisations, justifications, and perceptions from various types of agents involved, including policymakers, social workers, parents, and more. Local policy choices determine what is considered a priority beyond the legally defined criteria. Herman identifies two contrasting approaches: *formalism* in wealthier areas and *need-based* assignment in other areas.

The institutional setting, available information, and the complexity of forms and procedures have a significant impact on the smooth operation of markets (Pais, Pinter, and Veszteg 2011; Chen and He 2021). For parents, finding suitable childcare can be a costly and stressful process (Schüller and Steinberg 2022). The complexity of application processes may create optimisation frictions, influencing individuals' ability to make informed decisions, and either raising or lowering administrative barriers. Different information sets or media convey signals that can nudge some parents in or out of the applicant pool. In the childcare market, experiments show that detailed information and guidance and/or human support for applications have large positive impacts on registration and access for low-income families[15] (Weixler et al. 2020b; Zangger and Widmer 2020; Hermes et al. 2021, 2022).

[13] although before the law, 96% of children aged 3 where already in preschool.

[14] *article L. 214-7 du Code de l'action sociale et des familles.*

[15] This motivated another experimental research project joint with Carbuccia, Barone, et al. (2023), whose details are available on https://www.socialscienceregistry.org/trials/9901.

Priority scores and political signals In local administrations, time and efficiency constraints necessitate the use of simple and quick evaluation rules, typically in the form of priority scores. In the face of increasing demands, heterogeneous situations from registration to daycare entry, setting priority rules becomes complex. In 2018, the government assigned the French Mayor Association (AMF) to providing "*a framework for transparent and fair assignment of daycare seats*"[16]. The reports recommend to "*choose relevant priority criteria based on shared assessments, aligning them with the region's specifics*." Elected officials can endorse a charter to commit to these guidelines. However, they lack details on the allocation process, leaving implementation questions unanswered. The recommendations remain vague, with limited normative guidance or political accountability.

Most surprisingly, the assignment procedure *i.e. how* to assign daycare seats once priorities are set is barely mentioned in the AMF's report[17]. The only part that refers to assignment procedures is about computer tools[18]. And yet, it is well known that there can be many assignments that respect priorities with very different welfare implications. For instance, the student-proposing or college-proposing deferred acceptance algorithms (Gale and Shapley 1962) find stable assignments, but the first is the best outcome for all students and the latter is the worst. The first one is strategy-proof for students while the second is not. The choice of assignment method is crucial, and each method has its advantages and disadvantages. Transparency is a necessary condition for *fair* assignments but by no means a sufficient one.

Because there are no clear guidance on assignment mechanisms, priority scores appear as the only available tool to regulate daycare access. They further serve as political signals, informing voters about early childhood policy implementation through the prioritisation or deprioritisation of certain groups. Neimanns (2022) explores the political economy of social investment and analyses the link between childcare preferences and voting behaviour. Using survey data from in eight European countries, his main analysis models the link between childcare spending support and voting intentions by income, conditional on other measures of political values and a set of context fixed effects. He shows that the higher up the income distribution individuals are, the more tightly voting behaviours connect to preferences towards childcare. Therefore, although lower-income individuals might be the strongest supporters of additional public childcare spending, left-wing and right-wing political parties have incentives to target reforms or define priority rules favouring more affluent voters because those voters' preferences translate more directly into actual votes. Moreover, recent work on politicians' perception of voters' preferences show that most politicians strongly believe voters' preferences are more conservatives than they actually are (Pilet et al. 2023). This result is true across all political groups, and may further amplify the previous implications. Finally, subsidies depend on household incomes and richer parents pay larger fees, adding another incentive to favour more affluent families.

Priorities also signal to parents their chances of obtaining a seat and affect the composition of the demand (Ünver, Bircan, and Nicaise 2018). For many territories, they are of paramount importance, and some have taken great care in their definition. The design of a priority score depends on the selected characteristics and their associated weights. Families may have multiple combinations of these characteristics, making the political dimension of selecting these features and assigning their weights particularly salient. However, in other territories, the priority score is less explicit. It usually serves as an internal tool for implementing a policy, indicating the considered criteria but not how they are considered.

Ultimately the daycare assignment procedures lack transparency and means to ensure due process of all applicants. Despite the undeniable shortage of daycare slots, perceived opaque or convoluted assignment procedures can exacerbate the perception of a more severe scarcity. This can erode trust in the procedure – even if it is entirely regular – and can lead to dissatisfaction among parents, tension between parents and administrators, and unnecessary social and political costs.

[16] The mission led by Élisabeth Laithier (2018) suggests the establishment of committees for fair assessments, informing families about timelines, encouraging their participation, and providing tailored support.

[17] The dedicated section insists on a clear calendar, on organising assignment committees.

[18] Laithier (2018): Pages 10 and 11: "*The use of decision-support tools may therefore be appropriate, if it can leave room for human expertise. [...] However, the use of these tools would not be sufficient for the exhaustive allocation of daycare slots, insofar as they would not be able to grasp the specific nature of each situation, and could leave out families who do not meet the criteria*". Authors translation.

II.2 Market designs of childcare marketplaces

Economists tend to adopt fairly expansive definitions of what counts as a transaction and (thus) what constitutes a market. In general, they are the institutions that organise transactions. Markets can be designed and their features affect "who trades, what is traded, and the terms of those trades" (Li 2017a). A. E. Roth (2018) defines marketplaces as "*infrastructures, rules, and customs through which information is exchanged and transactions are made [that] can be relatively small parts of large markets. Participants may have large strategy sets, i.e., many options available to them beyond those available in any particular marketplace*".

Market Design combines theoretical and empirical methods to build, analyse, and enhance such institutions (A. E. Roth 2018). It has led to various applications and policy successes over the past two decades[19]. The daycare assignment problem is theoretically and empirically close to the school choice problem for which there are well defined solutions and many large-scale applications. Their success are tightly linked to the theoretical properties of assignments which in practice, convey useful normative values.

Market design offers normative criteria Li (2017a) discusses ethics in market design and argue that "*Market design needs value judgements: we seek to design markets that function well, not badly, and 'well' and 'badly' have normative content*". He advocates for what he calls *informed neutrality*, *i.e.* using criteria from the market design literature to formalise many small value judgements. Typically, a matching that respects priorities is said to be *envy-free* which is often interpreted as a *fair* assignment in the literature. *Envy-freeness* facilitates *transparency* because it makes it easier to justify each individual assignment following priorities. Mechanisms that make it a dominant strategy to reveal true preferences are called "*strategy-proof*", because "participants don't have to make strategic calculations about what others are doing, they just have to decide what they like" (A. E. Roth 2018). Strategy-proofness is desirable from a normative perspective because it "*levels the playing field*" by preventing strategic players from improving their assignment at the expense of non-strategic players (Pathak and Sonmez 2008). Yet, even when a mechanism is strategy-proof, it may not be *obviously strategy-proof* (Li 2017b). Information and transparency affect participants' strategies. Details on mechanisms that make strategy-proofness salient increase truth-telling (Guillen and Hakimov 2018) but perception of risk may lead some participants to still try to game the system although there is nothing to gain (Hassidim, Romm, and Shorrer 2018). Conversely, strategies may be set up outside of the marketplace or through other markets. When place-based priorities are in place, residential mobility can also be part of parents strategies (Bjerre-Nielsen et al. 2023).

These notions of fairness take the objective of the algorithm's owner or designer as a normative goal. That is, Mayors and early-childhood department owns the power of defining the objective function of the algorithm. In our paper, we call them *policymakers* and emphasise the attached responsibility coming with that power. For Kasy and Abebe (2021), fairness provides a framework to critique the unequal treatment of individuals i with the same *'merit'*, where merit is defined in terms of *social weight*. They discuss three limitations of fairness-based perspectives under a procedural notion of justice: "*they legitimise inequalities justified by merit, rather than questioning the status quo; they are narrowly bracketed and do not adequately engage with the impact of algorithms on pre-existing inequalities; and that they do not consider within-group inequalities, leading to intersectional concerns*".

Merit is an important *narrative* to justify inequalities and priorities but 1) views on *deservingness* vary and can be changed (Stantcheva 2021; Hvidberg, Kreiner, and Stantcheva 2023) and 2) this narrative often serves "*merit as reward*", while more economic definitions would emphasise *merit as effectiveness*" (Durlauf 2008; Sethi and Somanathan 2023). Eliaz and Spiegler (2020) develop a model for competitive narratives, which she defines as a "*causal model that maps actions into consequences, weaving a selection of other random variables into the story*". Policymakers' narrative can be seen as a causal model, and may have testable implications. In the case of algorithmic decision, Kasy and Abebe (2021) propose to use causal inference and social weights to formally assess the fairness, equality and balance of power of algorithmic decisions. We take a similar approach in section V when we analyse the effects of different elements of priorities on inequalities.

[19] e.g. assignment of resident doctors to hospitals (A. E. Roth 1984), the matching of children to schools (Abdulkadiroğlu and Sönmez 2003) and teachers to schools (Combe, Tercieux, and Terrier 2022), social housing (Waldinger 2021), Kidney exchange programs (Akbarpour et al. 2020) and so on.

Daycare assignment mechanisms borrowed from the school choice literature Childcare allocation mechanisms vary across countries, reflecting diverse approaches. When assignment are centralised, research has drawn inspiration from the school choice literature, applying related findings to daycare allocation. For instance, Carlsson and Thomsen (2015) simulate daycare assignment using the student-proposing deferred acceptance algorithm (SPDA) on German data (Gale and Shapley 1962). Similarly, Herzog and Klein (2018) consider three German daycare marketplaces which are decentralised and sometimes uncoordinated. They propose dynamic versions of the *immediate acceptance* mechanism (Also called the *Boston algorithm*). Reischmann, Klein, and Giegerich (2021) implemented dynamic versions of SPDA specifically designed for decentralised contexts, and provides empirical evidence from two cities. Veski et al. (2017) evaluate the welfare effect of using a centralised assignment mechanism using SPDA instead of the previous decentralised system in Harku, Estonia. Kennes, Monte, and Tumennasan (2014) analyse the daycare assignment mechanism in Aarhus, Denmark, which is common to most Danish municipalities, including Copenhagen. The market is such that the oldest unassigned child is given high priority in a daycare where no current capacity restriction exists - a concept called “child care guarantee”. Children who are already assigned have the highest priority in those places in the subsequent period. Apart from age, some characteristics give higher priorities, such as children with special needs, siblings in the same daycare, immigrant children in need of special assistance in daycare.

Some settings use lottery-based assignment procedures which researchers leveraged as natural experiments to evaluate childcare policies and allocation processes. For example, Rio de Janeiro reformed its daycare market in 2008, introducing a lottery system to enhance transparency and O. Attanasio (2022) shows this shift positively influenced labour earnings and child development. Similar lotteries have been employed in Oslo, Norway, involving randomised slot assignments (Drange and Havnes 2015), and in Bologna, Italy, where applicants express preferences and are allocated slots based on a Family Affluence Index (Ichino, Fort, and Zanella 2019). However, being focused on evaluation research questions, these research provide little details on the assignment procedures.

New approaches to centralised assignment Kamada and Kojima (2023)[20] propose a theoretical framework for matching markets under multidimensional constraints. In their study, they employed data from two municipalities in Japan to simulate various scenarios to assign childcare with their algorithms. Market designs are not only applied to school choices and sometimes research on a different topic can be used to solve problem in others. In this case, we found a clear analogy between the recent work by Delacrétaz, Kominers, and Teytelboym (2023) on refugee resettlement and the daycare assignment problem. Indeed, both involve navigating multidimensional constraints related to family characteristics and needs (such as the number and age of children), the availability of suitable facilities across different locations, and the establishment of fair prioritisation criteria. We discuss both models in details in section IV and build our main results from theirs.

[20] At the time we began, a working paper was available at Kamada and Kojima (2019).

III A multi-site market design experiment

The following section presents the setting of our ongoing experiment and the data we use in the empirical framework. The project started in spring 2019 where we first actively engaged with local representatives and childcare service providers to gain a comprehensive understanding of how daycare assignment processes were structured in diverse settings. We organised group meetings, visited about twenty city-halls, interviewed professionals at various levels and observed the proceedings of two assignment committees. We started assigning daycare seats for the first time in April 2020 and recruited new cities over the years, guided by sample size requirements for causal evaluation of heterogeneous treatment effects. The project is ongoing and because of that, our data access are provisionally limited (See details in Appendix B).

In this paper, our analyses focus on two large urban centres, chosen for two reasons: first, one consider preferences over week-days, while the other does not. Second, their data and characteristics are well-fitted to analyse the effects of the main features of these markets: the roles of priorities and diversity requirement in daycare accessibility inequalities.

III.1 Case Study I: A large marketplace with demands over weekdays and diversity constraints

Most of the analysis focuses on data from Valence Romans Agglo (VRA or Case Study I hereafter), an urban community in the Drôme department, southeastern France. It involves 54 municipalities and 223,630 inhabitants in 2020, half of which living in the two main cities[21].

A) Context

This marketplace provides access to 33 collective daycare centres, known as "*multi-accueils*". Families can apply through two central registration points (*guichets uniques*). Additionally, they propose two *crèches familiales* where accredited childminders provide care. Figure C.13 displays a map of the area, plots the daycares and the density of demands in 2023. It shows that this marketplace is very wide[22] and daycare centres are mostly located in the two largest cities and close suburbs. Most demands also come from these urban centres. Our collaboration with VRA spanned from spring 2020 to 2023, covering four years of operation.

In this setting, parents can register throughout the year, but the assignment committees are conducted in two rounds, typically in April and May for entries in September. At the end of March, the first round[23] of assignments takes place. A small allocation committee then reviews the proposals and makes any necessary adjustments. Families are offered a first-round assignment or placed on a waiting list. Assigned parents must promptly confirm their registration by scheduling appointments with their designated daycare centre. A second allocation round is held to fill any remaining vacancies, including those from the first round and those that became available due to withdrawals or changes in scheduling. This list includes families who were rejected in the first round, those who declined the first-round offer but remained on the waiting list, and new applicants. In September, a small committee meets to update any vacant placements and allocate them to families at the top of the waiting list, if possible. The specific procedures for these ad-hoc committees are not publicly available. Last, the early childhood department also has an emergency (temporary) process in place for parents who suddenly require childcare arrangements due to unforeseen circumstances such as divorce, violence, or job changes. Additionally, they offer occasional care services.

Our experiment started in 2020 and we ran assignments for four years. Obviously, 2020 was not the easiest year to start off with. First, there were local elections in March that prevented current incumbents to sign any contract over the next term. Then, the COVID-19 pandemic broke-out, delaying (among other things) the second ballot. However,

[21] https://www.insee.fr

[22] Going from the centre of Valence to the centre of Romans-Sur-Isère is a 30 min car trip or 40 min train.

[23] A small number of high-priority files do not participate in this marketplace. These files typically involve parents or children with specific needs, such as disabilities, social support or surveillance from child protection services, or underage parents. Social workers meet before the assignment committees to determine how to handle these special cases. Only after this pre-assignment round do we know the capacities to be assigned.

the *Communauté d'Agglomération* was not affected by the elections and maintained its commitment to the project just as the pandemic was beginning. Paradoxically, this crisis period really helped to prove the usefulness of our automated assignment as it freed civil servants in the early childhood department from this tedious task in a moment where their time was a scarce resource. Despite the pandemic *per se* and working remotely, they were in charge of organising childcare for *essential workers*, helping parents in distress, dealing with unemployed childcare workers and more. Furthermore, organising the assignment committee remotely was technically difficult at that time. This first success then convinced other municipalities to join the programme and the early childhood department provided positive testimonies to other municipalities.

B) Data and notable features

Data and main variables of interest We use two sets of data:

- Family datasets: they contain households and children IDs, reported preferences, priority score and sociodemographic variables used in its formula. These tables also contain results of the assignment, propensity scores and assignments from alternative simulations and approximate GPS coordinates[24].
- Supply datasets: They contain daycare ids and, number of seats for each age group over weekdays and age limits. It indicates which age group accepts children from the older or younger groups.

The family datasets contains one line per demand, and there are several demands for the same child and from the same family for different children in different assignment waves. We stack all family datasets and build our main variables of interest: age at registration, social groups and distances to reported daycares. The committee defines 5 social groups defined as dual earner couples, single earner couples, active single parents, inactive single parents and inactive couples. They give each group different weights and they can have other characteristics considered in other criteria. We do not have other socio-demographic variables. Since parents can register as soon as the pregnancy is official, child-age at registration is used as a proxy of anticipation and may be different across social groups.

We have several simulations of alternative assignments for 2022 and 2023. For 2022, we simulated assignments with priorities defined like in 2021, with the new priorities without the one for time since registration and fully random priorities. For 2023, we run similar simulations and also implement alternative mechanisms.

We also build a dataset of all ranked preferences and results of the assignment. In practice, this dataset allows to track each assignment and justify decisions individually. We use this dataset to showcase the properties of our proposed solutions.

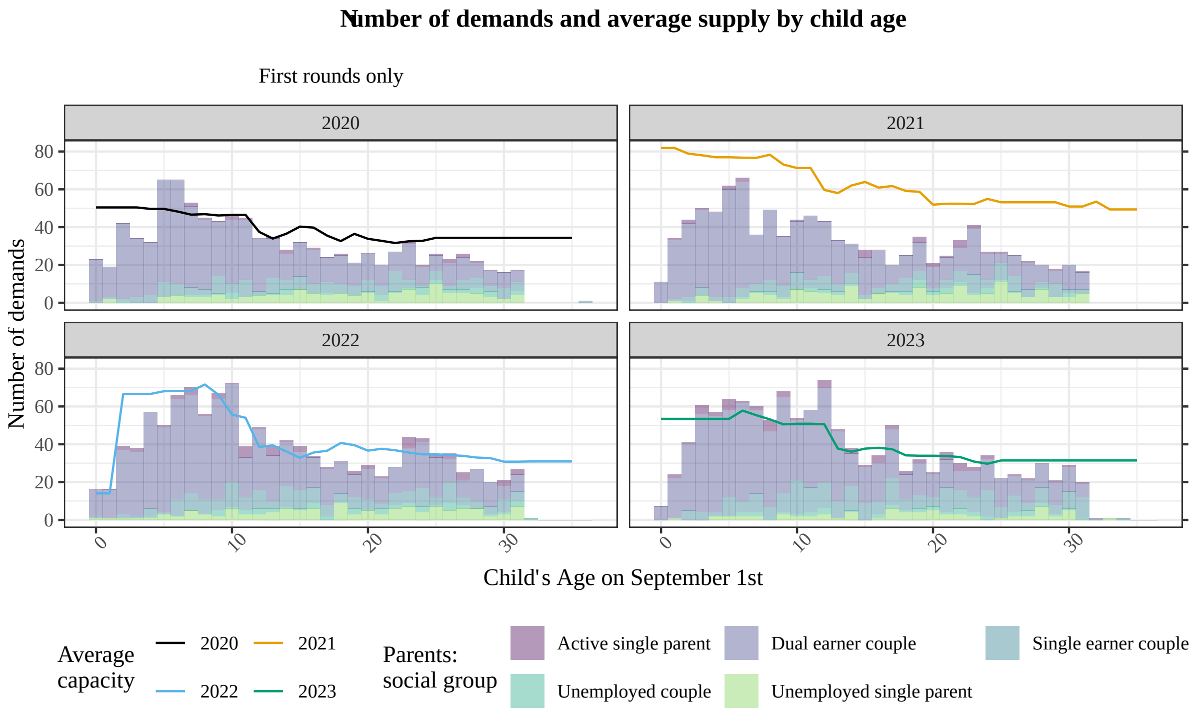
On the supply side: capacities by weekdays for different age groups As discussed in section II, daycares often have age groups with designated area and capacities which may vary from one daycare to another. In this setting, there are up to five age groups in a daycare and the supply datasets report capacities by age groups and over weekdays. Moreover, some age groups have flexible definition and can accept children from other groups, if there are enough capacity. The problem involve eligibility rules and a nested priority ordering for some buckets of capacities sorting files by age group then priorities.

Table C.3 in the Appendix summarises the aggregated results from various allocation committees. During the four spring committees, daycare centres offered an average equivalent of 415 placements for the first round. In this setting and like in all cities, daycares offer much more seats for younger children than older ones.

One way to apprehend the differences in access based on children's ages is to count, for each age value in September, how many buckets are open in all daycares and how many slots they offer on average over weekdays. By summing the number of slots in all buckets, we obtain the "potential" of open slots by age. By dividing the potential by the number of sections that open it, we obtain the average potential seats in each daycare for a child of a given age. Figure 1.1 presents these estimates along with an histogram of demands. On average, for first rounds, children have between 5 to 7 slots offered if they are under 12 months old, and around 3 or 4 if they are older.

[24] The GPS coordinates comes from the GoogleMaps API from a list of addresses and have been rounded for privacy purpose. We rounded latitude and longitude at the 100^{t}h degrees which creates points every 100 meters.

Figure 1.1: The share of open seats by age follows the distribution of age in the demand



Sources: ISAJE, Case Study I − 2020 : 2023 − first rounds only.
Note: − Histograms (left scale): number of demands by children age in September, by social group. Binwidth of 1 month.
− Solid lines (right scale): Ratio of the sum of the average number of days offered in all buckets accepting children of this age to the number of buckets opening a slot.

On the demand side: 1140 demands ranking up to 8 daycares reporting preferences over weekdays Table C.4 in the Appendix reports the number of demands for each round. In each year of the experiment, we processed an average of 1144 applications in the first rounds and 834 in the second rounds. Between 2020 and 2021, the demand increased by 5%, but it was in 2022 that the progression was the strongest: +17% compared to 2021. In 2023, there are 25% more requests than in the 2020 Committee. Table C.5 and Figure C.14 show the evolution of demands by social groups and how the proportion of each category evolves each year. Most demands come from dual earner couples while there are very few cases of single parents working. Over the years, the number and share of couples with only one parent working and single parents increased.

Families can rank up to 8 daycare centres, but only 1.1% of families provide a fully complete list. Figure C.18 in the Appendix shows that 1/4 of families reported only one choice and half request only 3. This means that parents' choices are not limited by the 8 preferences. Figure C.19 in the Appendix shows the average proportion of demands over weekdays for each year of assignment and by main social groups. Figure C.20 shows the cumulative distribution of the number of days reported. Overall, more than 50% of all demands register for 4 or 5 days, there are less demands for Wednesdays and significant heterogeneity across social groups and over the years.

We provide extra analysis of preferences in Appendix C.V with additional comments. Figure C.21 plots the *heat map* of the average number of applications over years for each daycare and preference order rank. The 5 largest daycare centres are also the 5 most demanded across all ranked choices. In the map presented in Figure C.22, we show the flows from parents' home towns to each requested daycare centres. The two urban centres serve as major hubs, concentrating nine of the most sought-after and, consequently, most congested daycare centres. Parents out of the two large urban centres tend to choose the closest daycare as first choice and then others in the main urban centres. In Figure C.23, we present the distribution of distances to various types of childcare centres. Registered parents live close to the most demanded daycare centres, especially the most congested ones. This is endogenous selection at play and we do not observe the distribution for parents who did not apply. We can only say that among those registered, most live very close to their chosen daycares.

Priority criteria and weights In this setting, the application forms and website do not provide parents with much information on how the assignment committee works. They encourage parents to "*contact early childhood services as soon as the pregnancy is official. Daycare seats are allocated according to availability and a number of criteria: residency, family situation (social or health-related), family structure and employment and time since registration. Fees depend on incomes and household composition*[25]". However, parents do not know how these criteria are weighted and used and children age is not reported. At best, it is implicitly implied that availability could depend on children age.

In practice, the early childhood department defines a weighted sum of 10 variables built from these criteria. In Appendix C.IV, we describe precisely how the score has been built and evolved. In brief, family situation defines 3 criteria: based on parent health, job search or training ; children health or disability and when child protection services are involved ; siblings - additional priority for multiple demands (e.g. twins) or when a sibling is already in one daycare. Family structure and employment define the 5 variables classifying social groups and strongly favour dual earner couples and single parents working. The last two variables are time since registration (in months) and living in the urban area.

The priority score is a constraint of the daycare assignment problem which guides us toward assignment free of *justified envy*. These criteria and weights have been voted by incumbents from the 54 municipalities and elected officials take pride in this consensus[26]. We asked elected incumbents and heads of early childhood department why these criteria and weights where chosen, and if they considered them "fair". First, territorial inequalities were central to the debates and they voted for equal priorities for all parents in the urban area. Second, they emphasise the balance between prioritising "*families with higher needs*" while rewarding parents who anticipate and those who wait longer. *Deservingness* justifies the weights, the residency bonus and time since registration. Anticipation and willingness to wait are perceived as a signal of higher needs and a mean for parents to increase their chance of receiving a seat by registering earlier than their competitors. We call *social weights* the part of the priority scores without time since registration as they reflect policymakers valuation of their social situation.

Parents do not know their social weights but know that registration is open as soon as the pregnancy is official and time since registration increase priorities. For parents, striving to be earlier than others can be a dominant strategy when the system is partly "first come, first served". The readier they are, the better, as registration can start on the third month of pregnancy. These rules can also affect the demand at the extensive margin $i.e.$ discourage applications from parents whose pregnancy is poorly timed with regards to the assignment schedules. At the intensive margin, densities of registration relative to the child birth can shed light on strategic behaviours.

The consequences of these strategic behaviours on the distribution of assignments depends on how much time since registration explains the variations in priorities. The remaining variance depends on joint-distribution of family situations and social groups. For the latter, Table C.6 in the Appendix presents the distribution of all criteria by social groups over the four years of the experiment. Family situations actually concern relatively few demands, although, in proportion, couples with no jobs are more likely to have children or parent priorities. It is striking to see that the social group with the highest priority – working single mothers – has very few demands although their number increases over the years. An important feature of this market is that dual earner couples represent 66% of the demand and only 14% of them have family situation priorities.

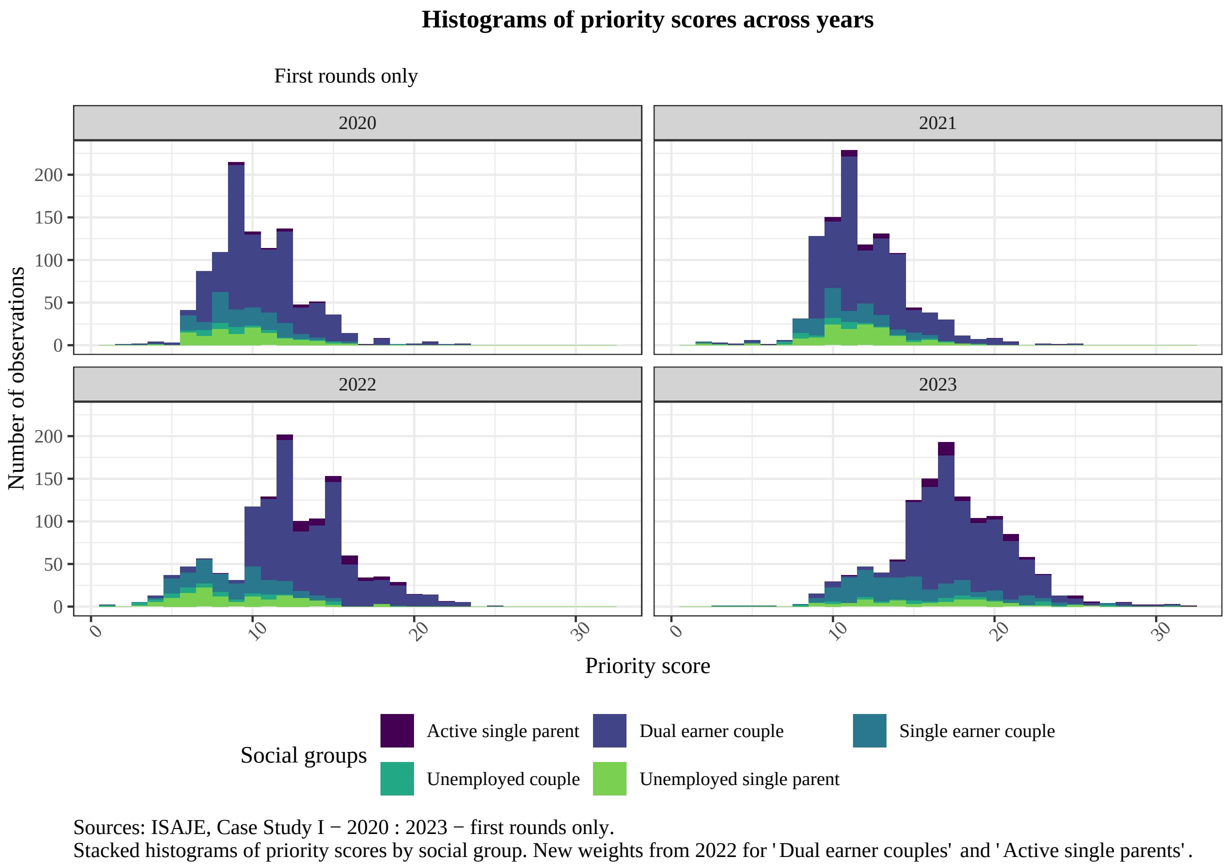
Figure C.16 in the Appendix presents the scatterplot between months since registration and the priority scores. We simply fit an OLS regression to estimate the share of variance explained. In 2020, 78% of the variability in the priority score come from time since registration. The criteria used are coarse and without time since registration, there would be few large groups of priority ties. Early registration is a strategic parameter that breaks ties within social priority groups Timing of registration is therefore a highly strategic parameter for parents, especially for dual earner couples. They represent the majority of demands and mostly compete against each other for early entry. Early registration is their main strategic lever to increase their chance of being seated.

[25] Own translation of the short description of assignment committees available on the city-hall website.

[26] The president, Nicolas Daragon, was first elected in 2016 and re-elected in 2020. Member of *Les Républicains*, the historical right-wing party, he has been the mayor of Valence since 2014.

The 2022 priority weigth reform In 2022, priorities were reformed, dramatically increasing weights for dual earner couples and active single parents. From 2022 onwards, weights for dual-earner couples have been multiplied by 2.6 going from 3 to 8 points, those of active single-parents by 2.25 going from 4 to 9. Inactive single parent families and couples now receive 2 or 3 points, while unemployed couples still receive 1 point. Figure 1.2 presents the distribution of scores for the first rounds of assignment for all four years, stacked by social group. In 2020 and 2021, the construction of scores is similar, and the two distributions are close. There are significantly more dual-earner couples than all other situations combined in all the allocation rounds. From 2022 onwards, densities shift for dual earner couples to higher priorities. In subsection V.1, we investigate the consequences of this change by simulating alternative assignments and measuring the effects on segregation.

Figure 1.2: Evolution of the distributions of priority scores over the years by social group



III.2 Case study II: A simple school choice problem with diversity constraints

The previous case study considers the daycare assignment problem with preferences and capacity over weekdays although only a quarter report less than 4 days. In other settings, policymakers opted for a simpler version of the problem considering only "*full-time*" demands in the assignment procedure. Demands for fewer days are only matched after the main assignment took place on the remaining capacities. In these settings, our experiment only concerned the full-time assignment problem, leaving the part-time market aside. Our second case study is one of those.

The main conceptual difference is that without preference over weekdays, one child takes one seat and the problem is well known. Our contract with this local administration does not allow to reveal much details and the main point of this second case study is to emphasise that i) policymakers choose the definition of the problem, ii) this affects the structure of the market and parents options iii) diversity constraints have more influence on assignment probabilities than priorities. In this setting they cause sharp inequalities by date of birth.

Context In this setting, parents can register from the 6^{th} month of pregnancy through a *rendez-vous* at the city-hall and rank up to 5 daycares for 4 or 5 days. Like in VRA, they are informed of the criteria used in the assignment process but do not know how. The main difference with VRA is that i) they do not use time since registration, but a small bonus for those who update their files on time, and families whose demands for another child have been rejected in previous years. ii) they have more options for various criteria, but the number of ties is high (Figure 1.3).

Daycares are divided into age groups with strict eligibility rules. In VRA, several buckets allowed children from other age groups, it is not the case here. Moreover, that year, the first age group has the same definition in all daycare: children born on that calendar year. Groups for older children may vary.

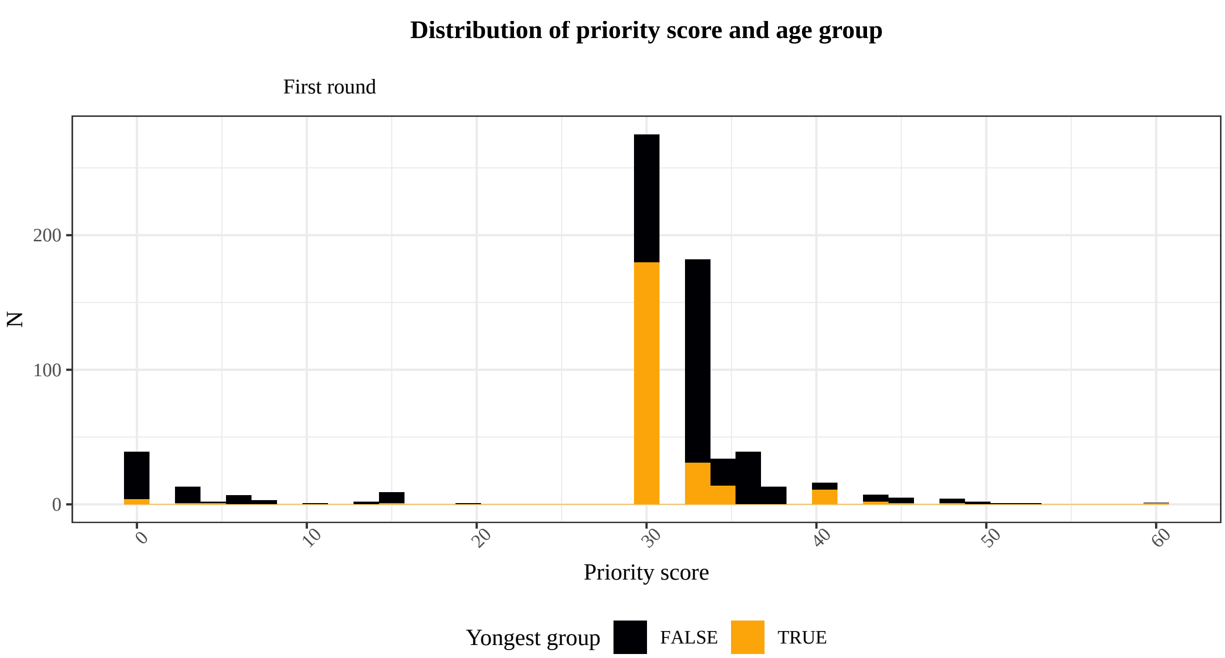
Data and descriptive statistics We use data from the *supply* file of case study II which contains 15 daycares, buckets birth date limits and capacity. The *Family* file of case study II contains 657 demands with family ids, date of registration and date of birth, ranked preferences priority scores and criteria. It also contains the results of the assignment, the lottery realisation and propensity scores.

Figure 1.3 shows the distribution of priorities colour-coded for children born that year and older ones. The priority score is essentially divided into three groups:

- Priority under 30 represents 12% of the demands. Their shared trait is that they don't work, and their children are not born that year ;
- Priority of 30 represents 42% of the demands and corresponds to the minimal score for parents who work. They are 3/4 with children born that year.
- Priority higher than 30 represents 46% of the demands and includes parents who work with extra criteria. Most of them were not born that year.

There are lots of ties for parents of the main group with no specific characteristics: dual earner couples, 3/4 of them with children born on the first semester of that year.

Figure 1.3: The most frequent priority is 30, those are mostly children born that year.



Sources: ISAJE − Case study II. Histogram of priority scores. Youngest group defined as born after January 1st that year.
Notes: Binwidth of 1.5. 12% of the demands are below 30, 40% of the demands have a priority of 30.

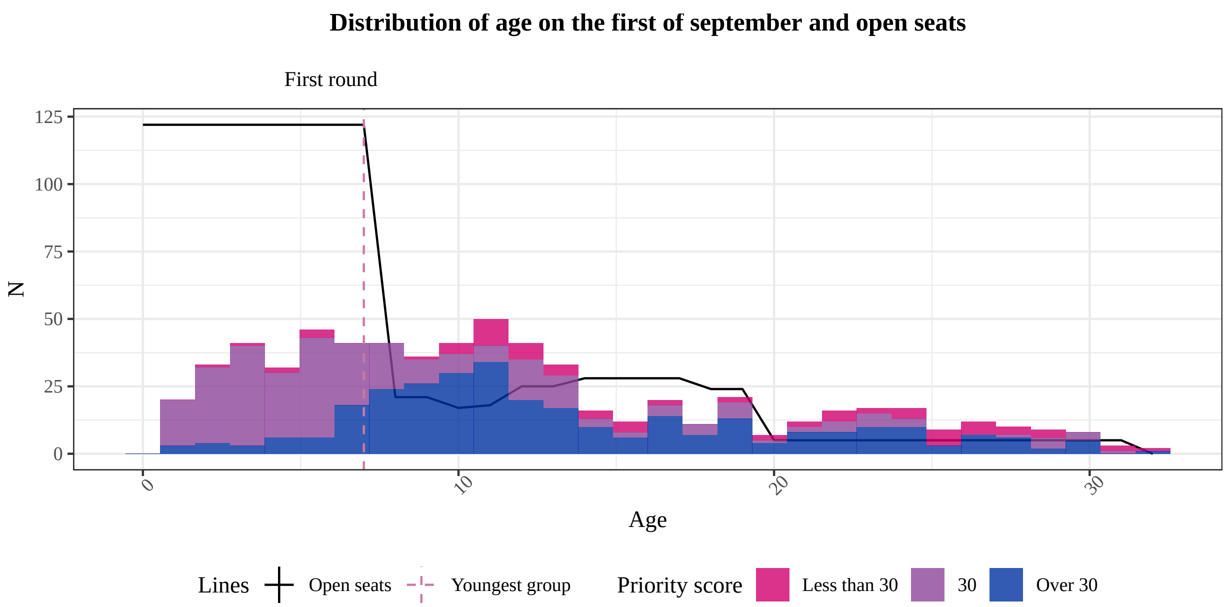
That year, daycares offer 165 seats among which 122 are available only for children born this calendar year (younger than 8 months in September). We compute the total number of open seats across daycares by children's age in

September. We plot the result in Figure 1.4 with the histogram of children's age in September by priority group. This figures shows that:

1) There is a sharp discontinuity in the number of offered seats for children older than 8 months in September but roughly the same number of demands on either side of the discontinuity ;
2) Eligible demands for the most available seats are mostly working parents with no other priorities ;
3) Children with priorities higher than 30 are mostly older than 8 months and have far fewer open seats.
4) Those with priority lower than 30 (who do not work) are also mostly over 8 months and compete with those with highest priority over few seats.

We come back to these two case studies in section V.

Figure 1.4: Young children get 100 more seats, demands and priority are evenly distributed across age limits



Sources: ISAJE − Case study II. Histogram of age on September 1st, stacked by priority groups.
Notes: 12% of the demands are below 30, 40% of the demands have a priority of 30.
Open seats are computed as the sum of slots across all buckets with open seats for children that age.
Youngest group indicates the common age threshold across daycares that define the youngest group.

So far, there are two versions of the daycare assignment problem. The difference is a political choice of simplification and sequencing. Case study II consider two separate markets: one for full-time and one for part-time. Full time demands is the main marketplaces and the others are used to fill the gap. Case study I consider demands and capacity constraints over weekdays and combine as many demands as possible while following priorities and capacity constraints. In practice, many rules stems from practical constraints and we adapt our algorithms and codes to local specifics. An important shared feature is that seats in daycare are reserved for specific age groups. This makes it hard for policymakers to justify assignment using the scores. Indeed, because of age groups, children with the same score asking the same daycare may have different outcomes because they are not eligible over the same bucket. Building from these features and constraints, we now move to our theoretical models.

IV Models for the daycare assignment problem

We consider any daycare marketplace where a local authority is in charge of assigning slots to families through a centralised assignment mechanism. For consistency with the school choice literature we call applicants "students" and daycare centres "schools" throughout the paper. Like school choice, application processes involve submitting an ordered list of **preferences** and reporting a set of characteristics used to sort students with **priority levels** and set eligibility status. For now, we consider individual applications[27]. Recall that our goal is not to provide new theories but adapt and possibly improve existing work so that they could be applied in the field. Our models are very general although they reflect our empirical settings. Like assignment committees, they i) take agents participating in the daycare marketplace and their characteristics as fixed, and ii) do not consider the problem of self selection, where preferences come from, and the role of outside options or dynamics.

The goal of this section is to provide a formal definition of **daycare assignment marketplaces**, and for that, we need to i) enumerate the **elements** of the problem, ii) introduce appropriate **stability** notions, iii) define **algorithms** to find stable matchings, if they exist and iv) enumerate their **properties**.

We opt for a pedagogical approach and start by introducing notations and take the student proposing deferred acceptance as a benchmark in subsection IV.1. Subsection IV.2 adapts the problem with diversity constraints and solve the simplified problem without demands over weekdays. In subsection IV.3, we consider the problem with demands over weekdays and we present our main result in subsection IV.4. Finally, we show the properties of our assignment in action and compare the outcomes of our various algorithms in subsection IV.5.

IV.1 Notations

We denote I the set of individuals, where $i \in I$ is a specific student. Student i is characterised by their **type** θ_i which captures all the relevant information for the assignment of individual i. We denote by Θ the set of possible types of each individual and by $\boldsymbol{\theta} := (\theta_i)_{i\in I} \in \Theta^N$ the vector of individuals' types. There are S schools indexed by integers $s = 0, 1, \ldots, S$ where $s = 0$ is the "*null school*", *i.e.*, being unassigned.

Each applicant $i \in I$ has a (strict) **preference order**[28] $\succ_i$ over the set of schools,[29] where $s \succ_i s'$ indicates that family i prefers school s over school s'. Applicant $i \in I$ in school $s \neq 0$ is assigned a **priority level**[30] $\rho_{is} \in \{-\infty, 0, 1, \ldots, K\}$. Priority levels are **given** by municipalities. They typically use scalars summing and weighting a (small) set of criteria. If $\rho_{is} > \rho_{i's}$, applicant i has a higher priority than applicant i' in school s. Priorities are usually coarse and there can be many applicants with the same priority applying for a school. If $\rho_{is} = -\infty$, applicant i is ineligible in school s. If $\rho_{is} = 0$, applicant i has no particular priority for the school s. A school s is said to be **acceptable** to family i if $s \succ_i \emptyset$.

A school s has q_s^t seats available at day t. We let $q_s = (q_s^t)_{t\in T}$ be the vector of **capacities** of school s and $\mathbf{q} := (q_s)_{s\in S}$ be the one for all schools. Schools have capacity and diversity constraints. They can come in many forms but we can define constraints very broadly. A **constraint** at school s is a non-empty collection $\mathcal{F}_s \subseteq 2^I$ of sets of families. We say that a subset $I' \subseteq I$ is **feasible** at s if $I' \in \mathcal{F}_s$ and it is **infeasible** otherwise.

In the matching literature, a policymaker typically uses a **matching mechanism** ϕ, also called **matching algorithm**. An algorithm simply maps each element of a problem P to a matching (μ) of these individuals to schools.

[27] Putting aside twins and families with multiple demands of children of different ages following current practices. Parents in such situations usually receive higher priorities.

[28] These preferences are lexical *i.e.* they describe preference *hierarchies* and the idea is that (under certain conditions), an agent's rankings of their possible choices can be represented as if they assigned a level of utility.

[29] Note that the schools contain the "null school" so that an applicant can prefer to be unassigned than being assigned to some schools.

[30] Our notations define priorities in a reversed way to Abdulkadiroglu et al. (2017) for whom a lower ρ gives a higher priority. We make this choice to better reflect the practices of municipalities that generally use a weigted combination of criteria or where a higher level gives a higher priority. In practice, we always redefine priorities with lexicographic notation to nest different priority levels. For instance, if there is a score defined in tens, we will use the hundreds to code the age priorities which are superimposed on the score within daycares.

Stochastic assignment as a research design Priority levels typically create *ties* among applicants at a given school. To select among such candidates, we use a **single random tie-breaker**. For each individual i, we draw a random number ϵ_i from the uniform distribution $\in [0,1]$ and define π_{is}, the **applicant score** at school s as a combination of priority level and random tie-break. For each realisation of the tie-breaker, an algorithm returns a matching. For a given problem, a mechanism generates a distribution of probabilities over possible matchings, which is referred to as a **stochastic assignment**. A stochastic assignment generates a matrix $\mathcal{P}$ of size $|I| \times S$ where the entry p_{is} represents the probability that applicant i is assigned to school s. $\mathcal{P}$ is key to later estimate the effects of accessing daycares or estimate compliers' characteristics. We literally embed randomised experiments at various steps of the algorithm and *harvest* these variations to identify causal effects of interest. In Appendix F.III, we provide a formal proof that our stochastic assignment methodology can be employed for estimating design-based propensity scores and later be used for causal evaluation.

The school choice problem: a brief summary In the so called *school choice problem*, demands and capacities are **unidimensional**, each school s can admit at most q_s students, and we denote $\mathbf{q} := (q_1, \ldots, q_S)$ as the vector of capacities. Applicant i's type is a combination of their preferences and priorities, that is, $\theta_i = (\succ_i, \rho_i)$ and Θ is just the product set of all possible preference profiles and vectors of priority levels. In this case, the policymaker faces the **assignment problem** $P_{1:1}(I, S, \mathbf{q}, \boldsymbol{\theta})$. The main features are preferences from parents and schools/policymakers through priorities and capacity constraints. As exposed in section II, this problem is well known in the school choice literature and already implemented for the daycare assignment problem in various settings. An important definition is the stability notion for this problem.

Definition 1.1 (Stability in school choice)**.**

Given a matching problem $P_{1:1}(I, S, \mathbf{q}, \theta)$, a matching μ is said to be **stable** if there exists no **blocking pair** (i, s) of an individual and a school s.t. $s \succ_i \mu_i$, so individual i also prefers school s and either:

- $|\mu_s| < q_s$ and $\rho_{is} > -\infty$: school s is not full and individual i is eligible to school s.
- $\exists\, i' \in \mu_s$ s.t. $\pi_{is} > \pi_{i's}$: school s is assigned another individual i' which belongs to an applicant with a lower score than i at s.

If such pair exists, some applicants will be able to "complain" about this assignment. If the complaint is that there are empty seats left in a school, we say that the matching is **wasteful**. If the complaint is due to the existence of an applicant with lower priority assigned to the school, we say that the matching has **justified envy**. In this setting, a stable matching is simply a matching that is **non-wasteful and envy-free**.

There may be many stable matchings. Among these, one can show that there exists one that is the most preferred for all applicants. This can be found using the *Student Proposing Deferred Acceptance* (SPDA) algorithm proposed by Gale and Shapley (1962) and defined in this setting in Algorithm 1 in the Appendix. The SPDA has many attractive theoretical properties. It is is the **only mechanism** that is **stable** *i.e.* non-wasteful and envy-free, and **strategy-proof** *i.e.* applicants have an incentive to truthfully report their preferences[31](Abdulkadiroğlu and Sönmez 2013).

Moreover, the SPDA algorithm can also be seen as a cut-off adjusting function[32] and Abdulkadiroglu et al. (2017) derive their propensity score formula from this definition in a continuum economy. At each round of provisional assignment, the rank (or priority) of the last provisionally accepted applicant defines an admission cut-off. At any round, applicants with a rank lower than the cut-off defined in the previous round are rejected. Applicants with a rank higher than the cut-off will increase the cut-off to the rank of the last feasible provisional assignment and so on.

Considering the two cases presented in section III, it is clear that this does not fit the first setting (VRA) because of preferences over days. However, Case study II is very similar and only adds age groups constraints. As we will see, the school choice problem with diversity constraints can be transformed into a $P_{1:1}$ problem and solved using SPDA.

[31] Precisely, it is a dominant strategy for them to be truthful: whatever the report of other applicants is, one cannot benefit of reporting a preference list other than their true preferences.

[32] See theorem 1 in Kamada and Kojima (2023) and the characterisation of the space of cut-off profiles as a finite lattice and the characterisation of *stable* matching as fixed-points.

IV.2 The problem with diversity constraints

School are often structured to welcome children of certain ages in specific areas which defines **diversity constraints** within schools in the form of **eligibility or priority rules** attached to subsets of capacities. Daycare providers define their capacities for different **groups** of children. Groups broadly define categories (low-income families, etc.) that the policymaker would like to balance in each assignment[33]. We introduce new elements that we call "**buckets**" of capacity to designate bundle of capacities within daycares with attached eligibility and ordering over groups. This term is also used by Abdulkadiroglu et al. (2017) to designate reserved seats in the Denver Public School assignment. In practice, buckets often refer to age sections, but we prefer this more general notion.

A) Additional notations with diversity constraints

There are two main differences with $P_{1:1}$:

1) Each student $i \in I$ has a *group* $g_i \in G$ where G is the set of all possible groups. We denote g and G' as respectively a generic element and a subset of G. Following our previous notations, groups are now part of a student's type $\theta_i = (\succ_i, (\rho_{is})_s, g_i)$;
2) Policymakers and/or daycare providers define **buckets of capacity** with **eligibility** and **priority rules** by groups. Eligibility rules mean that some capacities restrict access to students of groups g only. Priority rules mean that they will favour students of group g but other groups g' may be accepted too, but only if there are no student of group g that can be accommodated in these capacities. This defines two sorts of diversity constraints:
   - **Soft quotas** are seats who accept students of groups g first but other groups can be accepted if there is enough capacity to accommodate them.
   - **Hard quotas** only accept students of group g. The others are not eligible.

Let us formally define the notion of buckets of capacities.

Definition 1.2 (Buckets of capacities)**.**

There is a set B of **buckets**. A bucket $b \in B$ is characterised by elements $(s_b, \succeq_b, c_b)$ where:

- $s_b \in S$ is the school of bucket b. We let $B_s \subseteq B$ be the buckets which belong to school s.
- $\succeq_b$ is the weak priority ordering of bucket b over $G \cup \{\emptyset\}$ where $\emptyset$ will be used to define the eligibility of a group for bucket b.
- $c_b := (c_b^t)^t$ represents the capacities of bucket b for each day t.

For our daycare problem, a bucket b can typically correspond to an age group in a daycare and the ordering $\succeq_b$ defines which age groups are eligible and prioritised. To fix ideas, assume that groups $G := \mathbb{N}$ correspond to the age (in months) of a child. If a bucket b of daycare s represents the number of seats that are proposed to children from 3 to 9 months old, if there are seats left, we authorise children from 10 to 18 months old to apply, then $\succeq_b$ is such that: if $g, g' \in \{3, \dots, 9\}$ or $g, g' \in \{10, \dots, 18\}$, then $g \sim g' \succ_b \emptyset$. If $g \in \{3, \dots, 9\}$ and $g' \in \{10, \dots, 18\}$ then $g \succ_b g' \succ_b \emptyset$. All $g > 18$ are not eligible so $\emptyset \succ_b g$.

[33] Groups can also designate special needs or specific types of accommodations. For instance, few seats with longer hours for parents with late shifts, children with disabilities. Since 2019, there are additional *diversity* subsidies for daycares with a share of low income and disabled children higher than a threshold.

a) A redefinition of priorities Policymakers still use a priority score ρ_{is} to sort applicants according to their criteria. However, assignments must respect the diversity constraints first. This means that they do not simply sort students by priorities in a school but within buckets. Now, priorities are defined differently in buckets depending on groups.

To fix ideas, consider a school s with at least two buckets j and k in B_s with $b_s^j(g)$ and $b_s^k(g')$ ordering groups in each bucket. Consider two students i of group g and i' of group g' such that[34] $\rho_{is} > \rho_{i's}$; $\mu_{i'} = s$; $s \succ_i \mu_i$. In other words, two students of different groups where i prefers school s to their assignment and i' has a lower priority score $\rho_{i's}$ in that school than student i. Based on the priority score only, i' appears to have justified envy.

In practice, committees are uncomfortable with such situations. They cannot justify an assignment based on the priority score ρ_{is} only. Yet they sometimes make the definition of the priority score ρ_{is} an important political tool. Weights and characteristics used act as a political signal to parents. The latter hold assignment committees accountable for respecting these rules. Unfair assignments can be politically costly for incumbents. But transparency without a stability notion with diversity quota can backfire. As Li (2017a) writes, « *A policy-maker may be familiar with the details of their environment, and yet not know how to state their ethical requirements in precise terms.* ». This is where formal market models of assignment committees bring more than fast and convenient computer programmes.

If we can define stability notions with diversity constraints and algorithms that can find such stable matchings, policymakers will be able to justify assignments and ensure transparent and due process of all applications. To do that, we make clear the transformation of priorities following the ordering of groups within a bucket.

Definition 1.3 (Priorities within buckets)**.**

For each bucket b of school $s = s_b$ and student i, we construct $\tilde{\pi}_{ib}$ the priority of student i in bucket b as follows:

- If $\emptyset \succ_b g_i$, then $\tilde{\pi}_{ib} = -\infty$ so that i is not eligible to b.
- If $g_i \succ_b \emptyset$, then $\tilde{\pi}_{ib} = \sum_{g \in G: g_i \succ_b g} \max_i(\rho_{is} + 1) + \underbrace{\rho_{is} + \epsilon_i}_{\pi_{is}}$

In plain words, we consider students in the daycare they apply and redefine priorities within buckets starting with eligibility. When students are eligible, we add to the priority score ρ_{is} and tie-break ϵ_i the **highest priority score** $\rho_{is} + 1$ **as many times as there are groups ranked lower** than the group g_i in that bucket *i.e.* $g_i(b_s)$. There are other functions we could have used. They only need to preserve the **lexicographic nested structure of priorities** with diversity constraints.

b) Mapping student preferences to buckets Finally, students report their preferences $\succ_i$ over schools but they will be assigned to buckets. Thus, to run an algorithm which matches students to buckets, we need to map preferences over schools into preferences over buckets.

Definition 1.4 (Preference order over buckets)**.**

For each school s, fix an arbitrary ordering $\gg_s$ over buckets in B_s. For each student i, we define $\tilde{\succ}_i$ as an ordering over $B \cup \{\emptyset\}$ as follows: for $b, b' \in B$:

- if $s_b \succ_i s_{b'}$, then let $b \tilde{\succ}_i b'$
- if $s_b = s_{b'}$ and $b \gg_s b'$, then let $b \tilde{\succ}_i b'$

In plain words, the preferences over buckets follow an arbitrary order within schools but otherwise respect preferences over schools. The only role of the ordering over groups is to tie-break how students apply over different buckets in a given school[35].

[34] We neglect tie-break here but in practice, stability is enforced over π_{is}.

[35] In the literature, this ordering is referred to as a *precedence order*. This ordering could have an impact on the final assignment. However, it is unclear whether one should follow an alternative definition of ordering. In our empirical work, we simply follow the order provided by municipalities, usually starting from lower age groups to older.

The previous definitions allowed to introduce buckets, quotas and new objects for priorities and preferences over buckets and we can define the general problem of the policymaker:

Definition 1.5 (The problem with diversity constraints)**.** With soft quotas, policymakers face a problem composed of **Students** I, **Schools** S **Buckets** B and their elements, **Capacities q**, **Student types θ**. Formally, $P_{DC}(I, S, B, \mathbf{q}, \boldsymbol{\theta})$

Intuitively, the careful redefinition of priorities and preferences transform the sophisticated problem with diversity constraints into a problem where students apply to buckets and each bucket is reinterpreted as a *tiny school* with its own set of capacities, eligibility and priority rules. Student types **θ** may be extended to account for preferences over days and capacities **q** be defined over weekdays.

Our strategy is simple, we take a problem like P_{DC}, apply definitions 1.3 and 1.4 to create new priorities and preferences with buckets. We then use well defined algorithm on the modified problem, retrieve this assignment and map assignment in buckets back into schools capacities. We start with the simplified problem where policymakers only consider full-time demands like in Case Study II.

B) Stable matchings for the school choice problem with diversity constraints

We denote $P_{s:1}(I, S, B, \mathbf{q}, \theta)$ the school choice problem with diversity constraints. Again, this problem is well defined in the school choice literature and we adapt the work of Ehlers et al. (2014) to our setting. For a school s, a group g and a matching μ, we denote by μ_s^g the set of students of group g assigned to s under matching μ. Each school has a capacity q_s divided into quotas $\underline{q_s^g}$ and our *trick* is to define buckets b_s as **unitary seats**. We create b_s buckets and for each group g, we take $\underline{q}_s^g$ buckets with priority ordering $g \succeq_b g' \cup \emptyset$. In words, there are now **quotas** for groups g ; hard quotas only accept students of group g, soft quotas may weakly or strongly prefer students g over g'. Let B_s^g the set of buckets in school s reserved for group g. We have $|B_s^g| = \underline{q}_s^g$ and for $b \in B_s^g$, $g \succ_b \emptyset \succ_b g'$ for $g' \neq g$.

In plain words, there are individual seats. Each has their own priority order and a seat gives a priority to a given group of students. We now give a definition of stability in this version of the problem. The intuition is the same as before: we want to eliminate potential cases where students can have justifiable envy. We focus on the case with **soft quotas** $i.e.$ individual buckets with ordering over groups. We briefly discuss the case when all buckets are hard quotas with the properties.

Definition 1.6 (Stability with soft quotas in the problem without days)**.** For a matching μ, we let

$$\mu_s^g := \{i : \mu_i = s \text{ and } g_i = g\}$$

We say that a matching μ is **stable with soft quotas** if for any applicant $i \in I$ with $g := g_i$ and any school s s.t. $s \succ_i \mu_i$ either:

1. $\rho_{is} = -\infty$: *Applicant* i *is unacceptable at school* s, **or**
2. $|\mu_s^g| \geq \underline{q}_s^g$: *Quota is filled* **and**
   (a) $\pi_{i's} > \pi_{is} \forall i' \in \mu_s$ with $g_{i'} = g_i$
   (b) $\pi_{i's} > \pi_{is} \forall i' \in \mu_s \quad s.t. \quad g_{i'} := g' \quad \textbf{and} \quad |\mu_s^{g'}| > \underline{q}_s^{g'}$

In the case without quotas, we only have to respect the priority levels of applicants and fill each school. Now, with soft quotas, the only reason to violate the priority level of an applicant at a school is to respect the diversity quota at that school. This is exactly what Condition 2 of the definition imposes.

Let us detail all the cases hidden in Condition 2. If applicant i from group g prefers school s to their assignment, then:

- if the diversity quota of school s for group g is not reached, a student of group g would be able to complain. This is excluded by the condition $|\mu_s^g| \geq \underline{q}_s^g$ in the definition.
- If there is a student assigned to school s who belongs to the group of student g and who would have a worse priority at that school than student i, then the latter would be able to complain. This is excluded by condition 2a of the definition (note that the higher the value of π_{is} the higher the priority).
- Last, if we assign strictly more students at school s from another group g' than the objective of the quota and that one of these students has a worse score than student i, then the latter would be able to complain. Indeed, there would be no reason to give the seat to this student with a score worse than i because the diversity objective is already met. This case is ruled out by condition 2b of the definition.

SPDA with diversity constraints Again, many matchings can be stable with soft quotas. We use a generalisation of the SPDA algorithm proposed by Ehlers et al. (2014) and presented in Algorithm 4 in the Appendix. It simply redefines the problem into a 1:1 marriage problem *à la* Gale and Shapley (1962) by transforming preferences and priorities over and within buckets, then runs SPDA on the modified version. An important theorem in school choice is that "*a college admissions problem is stable if and only if the corresponding matching of its related marriage problem is stable*" (Abdulkadiroğlu and Sönmez 2013). Ehlers et al. (2014) show that this algorithm is **strategy-proof with soft quotas** and reaches the **student optimal stable matching with soft quotas**[36].

When policymakers define the problem with soft quotas and without days, we can solve and return the only assignment that is the most preferred by all families that strictly respects priorities and diversity constraints. In theory, this algorithm is strategy-proof so parents have an incentive to reveal their true preferences. The definition of buckets with group rankings allows to accommodate sophisticated constraints (through eligibility rules and local priorities over buckets) and to define target distributions over groups g in each daycare. Stability means that the assignment is envy-free and non-wasteful with soft quotas. However, if the definition of buckets is too narrow, assignments may be *artificially* stable and leave empty seats nonetheless. This happened many times in practice and while some adjusted buckets for re-running the assignment, others stood by the initial definition accepting empty seats and fewer satisfied family. Buckets are highly sensitive parameters and we think policymakers should consider them more strategically.

From a normative standpoint, stability ensures *fairness* and *transparency* by allowing individual justification of every assignment. It is also *safe* to participate in the market and reveal true preferences since there is no gain in trying to game the algorithm. We used this algorithm for Case Study II and other municipalities. As practitioners, we know this definition is a simplification of the *real problem* where parents' needs and daycares' capacities vary with days. The solution focus on a constrained market where parents are only eligible if they require full-time access, and part-time demands are considered after. This affects the composition of the demand at the extensive margin or may create other strategic behaviours. If the number of days reported is not binding, some could fill full-time demands and renegotiate their contracts once they receive an offer. At the same time, it can be hard for parents to anticipate their true needs so enforcing initial demands could make participation too risky for some. Conversely, using an assignment that directly accounts for preferences and capacities over weekdays could solve part of these problems. We now move to models with diversity quotas and multidimensional constraints.

[36] Ehlers et al. (2014) work with more general constraints than ours. They allow for upper and lower quotas. The algorithm presented here can be shown to be a special case when the upper and lower quotas coincide.

IV.3 Models with multidimensional constraints

In the situation without specific days required, one child takes one seat. A seat is an indivisible and unitary good. In settings with specific days, a student can take some days and the rest may be assigned to another with complementary preferences. A seat is therefore no longer indivisible and unitary. What changes is that childcare providers submit their capacities by days and parents rank daycares and their expected needs over weekdays. While a slot may represent a physical bound in a daycare (literally a cradle for each child), they are not tied to a particular child and different children can be assigned (and use) the slot in different days. Then, the allocation committees set priorities over applicants and consider combining the days of different seats to accommodate as many children as possible. The more general problem also involves diversity constraints and for that, our solution is to use the explicit redefinition of priorities and preferences using buckets to map the problem into a version with well defined solutions. In our case, the latter stem from the refugee matching framework of Delacrétaz, Kominers, and Teytelboym (2023) and of the constrained matching framework of Kamada and Kojima (2023)[37]. We start by introducing the new elements of the problem and a first important fact in problems with multidimensional constraint.

A) New elements in the problem with preferences over weekdays

The problem involve new elements in student types θ to map preferences over days, and changes in the definition of capacities. Each student now requires a set of days d_i. Assume that there are T days and that $d_i \in \{0,1\}^T$ where $d_i(t) = 1$ if applicant i needs to be assigned to a school day t.[38] We denote by $d = (d_i)_{i \in I}$ the vector of the **days required** by the applicants[39].

A school s now reports q_s^t seats available at day t. We let $q_s = (q_s^t)_{t \in T}$ be the vector of capacities of school s and $\mathbf{q} := (q_s)_{s \in S}$ be the one for all schools. In this framework, a matching μ is **feasible** if $\forall s \in S$ and $\forall t \in T$, $\sum\limits_{i \in \mu_s} d_i^t \leq q_s^t$. Equivalently, $\mu_s \in \mathcal{F}_s \ \forall \ s$. In plain words, a matching is feasible if for each day, the total number of assigned students to school s who required that day does not exceed the capacity of that day.

Recall that in the model without demands over weekdays, the stability notion is based on envy-freeness and non-wastefulness. However, as noted by Delacrétaz, Kominers, and Teytelboym (2023) and Kamada and Kojima (2023), it is easy to see that such stable matching might not exist in a setting with days.

Example 1.1 (Non existence of a stable matching)**.**

Assume there are 2 applicants i, i', one school s and that $\rho_{is} > \rho_{i's}$. Assume only 2 days ($T = 2$) and that s only has one seat available at day $t = 1$ and zero seats at day $t = 2$. Say i requires to be assigned both days but i' only requires to be assigned at day $t = 1$. In this example, there are no stable matching because

- There is no feasible matching that matches i to s since there is no seat at day $t = 2$.
- Matching i' to s would lead i to envy i' since i' has a lower priority at s.
- Leaving s empty is also not an option since, in that case, the matching would be wasteful since one can feasibly match i' to s.

A stronger result of both Kamada and Kojima (2023) and Delacrétaz, Kominers, and Teytelboym (2023) is that there are **no algorithm** that is both **envy-free** and **non-wastefull**. However, it is easy to show that an envy-free (but potentially wasteful) matching always exists. For instance, the trivial matching where nobody is matched to any

[37] In the terminology of the former, each day would be a *service* and our model would be a special case since we assume that each applicant requires only one seat per day whereas in the framework of Delacrétaz, Kominers, and Teytelboym (2023), a refugee family can require more than one unit of a service.

[38] In the daycare assignment of many municipalities, the days are typically the 5 work days, i.e. Monday to Friday. If one allows for half-days occupation then one can define $T = 10$ and interpret a given t has a half day.

[39] Here, we assume that preferences over schools and days are separated objects. A more general model would require to have preferences over the pairs schools and days assigned. We chose not to model it that way even though it is with loss of generality. First, many municipalities ask the parents to report exactly as in our model so that we chose to match their current practice. Second, negative results are already present and would probably be even more important in a more general model.

school is envy-free but obviously wasteful. The question is whether we can find matchings that are envy-free and not *too* wasteful. One can think that the condition of *envy* that we require may be strong. Indeed, i can block with s against a lower ranked applicant i' even though it would be impossible to feasibly match i to s while removing i' since they may require different days. This is indeed an important discussion of the literature of constrained matchings[40]. However, many different "natural" notions also fail to exist. The two aforementioned papers propose different notions of envy-freeness with different algorithm.

B) Chosing the right fairness notion

To understand the main issue with the definition of envy-freeness and the differences between Kamada and Kojima (2023) and Delacrétaz, Kominers, and Teytelboym (2023), we start with a simple example.

Example 1.2 (Reducing wastefulness)**.** In a daycare, there are four open slots every day except on Tuesdays. Say the family with highest priority asks all 5 weekdays and the next three families request different days than Tuesday (for instance, four days excluding Tuesday). The first family cannot be accepted because there is no open slots on Tuesdays, but the three others could. If we abide priorities, no family can be accepted and this daycare remains empty. This situation is bad for daycare providers and rejected families so policymakers may want to accept the three feasible families despite causing justified envy to the first.

a) Chosing weak-envy freeness Recall that the definition of a feasible matching in the multidimensional case is that $\forall s \in S$ and $\forall t \in T$, $\sum\limits_{i \in \mu_s} d_i^t \leq q_s^t$ i.e. there are enough capacities in each day in each school. Delacrétaz, Kominers, and Teytelboym (2023) introduce the notion of **weak accomodation** of applicants:

Definition 1.7 (Weak accomodation)**.**

School s can **weakly accommodate** applicant $i \in I$ alongside families $F \subseteq I \setminus i$ if $\forall t \in T$, either:

1. $d_i(t) = 0$ or
2. $d_i(t) + \sum_{f \in F} d_f(t) \leq q_s^t$

In plain words, this definition relaxes the original concept of feasibility by only taking into account the days in which i will take at least one unit of capacity. It has a **cumulative feature** considering every students that ever applied. If the aggregated sizes of applicants in F exceed the capacity of daycare s for some days, i cannot be accommodated alongside F; however, it may still be possible to weakly accommodate i if i 's size in that dimension is zero. In the example above, the three families can be weakly accommodated in this daycare as $d_i(t) = 0$ when t is Tuesday and $d_i(t) + \sum_{f \in F} d_f(t) \leq q_s^t$ for all other days.

By contrast, they define **strong justified envy** as a matching that allows **weak accomodation** but otherwise, respects priorities. Formally:

Definition 1.8 (Strong justified envy)**.**

Given a matching μ, applicant i in I has **strong justified envy** over family $i' \neq i$ with assignment $\mu_{i'} = s$ such that:

1. $\mu_{i'} \succ_i \mu_i$: student i prefers the school obtained by i'.
2. $\pi_{is} \geq 0$: i is eligible for seats in s and
3. $\pi_{i's}^s < \pi_{is}$ and
4. μ_s cannot weakly accommodate i' alongside all families with higher priority than i' at s that weakly prefer μ_s to their current match.

[40] We refer the reader to the discussion of Kamada and Kojima (2023) and other related papers of the two authors.

In this definition, we don't want a student i to complain that they are not in a school s they prefer to their current assignment μ_i when there exists student j with $\mu_j = s$ with $\pi_{js} < \pi_{is}$ i.e. lower priority[41] than student i in that school over days that *would have allowed i to be accepted*. Weak justified envy is the case where there exists student i' with $\mu_{i'} = s$ and $\pi_{i's} < \pi_{is}$ but there is no conflict over days between i and i'. Then we can define a matching that eliminates strong justified envy:

Definition 1.9 (Weakly envy-free assignment)**.** A matching μ is **weakly envy-free** if no family strongly envies another family.

Weak justified envy corresponds to the situation where we tolerate justified envy only towards those weakly accommodated. Weak envy-freeness relaxes envy-freeness by allowing some arguably innocuous priority violations.

b) Chosing strong envy-freeness Example 1.2 suggests that weak envy-freeness is a rather efficient way to limit wastefulness, with small deviations from priorities. The idea is that we can explain to such parents that accommodating these other students did not prevent them from receiving a seat. In some situations, policymakers may rather have larger vacancies than to deviate from priorities. There can be many reason for that, from transparency and accountability motives to organisational choices to offer occasional care on these remaining vacancies. It is therefore a **political choice** which affects the definition of a *fair* assignment.

In our setting, Kamada and Kojima (2023)[42] showed that it is possible to define matchings that are *student optimal* and *envy-free*.

Definition 1.10 (Student optimal fair matching)**.**

A matching μ is the **student-optimal fair matching** (SOFM) if :

1. μ is feasible, individually rational, eliminate justified envy **and**
2. $\mu_i \succeq \mu'_i \forall i \in I$ and every μ' that is feasible, individually rational, and fair.

Where **justified envy** is defined as any violation of priorities with tie-break. Formally:

Definition 1.11 (Justified envy)**.** An applicant i of group g in a matching μ has **justified envy** if $\exists s$ s.t. $s \succ_i \mu_i$ and $\exists i'$ of group $g \in \mu_s$ with $\pi_{i's} < \pi_{is}$.

In other words, definition 1.10 simply relaxes the *non-wastefulness* requirement, strongly enforces priorities and characterises the most preferred assignment. The notion of *optimal* matching does not mean that there is no other assignment that families would prefer to μ but that among those that eliminate justified envy, μ is the only one that is weakly preferred by all parents. Note that this definition includes student groups in order to be used with diversity constraints. In the setting without buckets, groups are irrelevant or equivalently, all demands are of the same group.

C) Algorithm for student optimal envy-free assignment

Together with the previous definition, Kamada and Kojima (2023) propose an algorithm - the *Cutoff Adjustment Mechanism* (CAM) - that always returns the SOFM. We formally present the CAM in Algorithm 2 in the Appendix. Intuitively, every student applies to their favourite school for which they have not been rejected yet. In each school, we sort the demand by rank $\tilde{r}_{is}$. The main *engine* is an **adaptive cut-off function** that checks feasibility in each school in every steps. At each round, every school that does not respect the feasibility constraint increases the cut-off by one, thus excluding a student that cannot be accommodated. At one point, the constraint is satisfied in every school and the assignment is final.

[41] with tie-break

[42] They consider a much more general setting than ours. They give a sufficient condition on their feasibility condition such that an optimal envy-free matching always exists. One can check that their condition is respected in our framework. We again refer the reader to Kamada and Kojima (2023).

Kamada and Kojima (2023) show that this algorithm stops in a finite number of steps and returns the **unique student-optimal envy-free** matching[43].

As we mentioned earlier, the matching returned by the CAM algorithm can be wasteful, $i.e.$ there might be an applicant who prefers a school which still has empty seats that would allow this applicant to be feasibly matched to that school (see example 1.2 above). However, by doing so, it would violate the priority of another applicant, who also prefers that school (which would in turn violate envy-freeness). The impossibility does not say whether the wastefulness "is large" or not, in particular it might be the case that, in practice, the amount of empty seats left is relatively small.

D) Algorithm for student-optimal weakly envy free assignments

Together with the notion of weak-envy freeness, Delacrétaz, Kominers, and Teytelboym (2023) propose the *Knapsack deferred acceptance* (KDA) mechanism[44] for refugee resettlement which we adapt to our setting in algorithm 3 in the Appendix.

This algorithm returns the **unique family-optimal weakly envy-free matching** (SOWFM). The main difference with the CAM algorithm and resulting matching lies in the interpretation of the concept of justified envy. The CAM procedure provides assignments that strictly adhere to priorities, even if it means leaving a potentially high number of seats vacant. KDA relaxes the definition of justified envy to accommodate more families as long as the total demand does not exceed capacity in any day.

Intuitively, this algorithm has a "cumulative" feature, that is, when a school temporarily keeps students, it considers all the students who have ever applied to it, even if they were rejected in an earlier step. It will weakly accommodate students up to capacity and adjust the cut-off based on the definition of weak accommodation.

How many more families can be accommodated with KDA rather than CAM depends on the setting. Because of the iterative and dynamic nature of the algorithms, a small change in some daycare can induce large changes in the pool of assigned families. Ultimately, both algorithms cannot ensure a non-wasteful assignment and our next question is: can we increase the number of assigned families further while ensuring a fair matching ? By optimality of both algorithms, it is clear that we cannot do better with these definitions of fairness. But we can further adjust the definition and see if we can improve welfare by sitting more students without harm.

E) No strategy-proofness in the assignment over weekdays

As in our previous section, one may wonder if the applicants have an incentive to truthfully report their preferences[45] $(\succ_i)_i$ in problems with multidimensional constraints. However, Kamada and Kojima (2023) show that **no algorithm is optimal envy-free and strategy-proof** and Delacrétaz, Kominers, and Teytelboym (2023) further prove that there are **no strategy-proof and family-optimal weakly envy-free** mechanism. They also propose an alternative algorithm that is weakly envy-free and strategy-proof but does not yield the most preferred assignment. We recently adapted this algorithm to this setting but did not use it in practice.

This negative result is strong: no algorithm can give the applicants an incentive to report truthfully while always returning an optimal envy-free matching[46]. However, it "just" means that we can find one example where one family, at a particular preference profile, knowing the preference profile of others, has an incentive to misreport. The theorem does not say whether "it is easy" to manipulate for a family, whether it is always possible to manipulate and so on. For instance, in large markets, the cut-off of a school is determined by the highest-priority applicant who

[43] In Kamada and Kojima (2023) terms, our setting respects the *general upper bound* property. Thus, the algorithm returns an optimal envy-free matching.

[44] In previous versions of their work, this algorithm was named *Cascading Multidimensional Deferred Acceptance* (CMDA).

[45] Note that here, we assume that applicants are only able to misreport their preferences over schools and not their required days. Indeed, it is consistent with our primitives since we assumed that applicants have separate preferences that do not vary depending on the days assigned. One can see the days as a hard constraint for families, say their job or obligations that they cannot change. Relaxing this assumption would only strengthen the negative result on the incentive properties.

[46] Note that the optimality is needed. Indeed, the mechanism which leaves all schools empty is trivially strategy-proof and envy-free. Whether there exists a non trivial strategy-proof and envy-free matching is an open question. Delacrétaz, Kominers, and Teytelboym (2023) propose such algorithm for a weaker notion of stability, we refer the reader to their article for details.

is rejected from it, and this depends on the entire distribution of student preferences as well as school priorities. It appears unlikely that any one particular student is in a position to influence the cut-off in any significant manner. Thus, one conjecture could be that such manipulations are rare in practice or too complicated for an applicant to find but more research are needed.

IV.4 Main results

A) A Pareto improving minor adjustment

By ensuring envy-free or weakly envy-free matchings, CAM and KDA provide assignments that can be wasteful. That is the price of envy-freeness. Kamada and Kojima (2023) tried to define *feasible envy* as a criteria for stable matching $i.e.$ considering that those with infeasible demands cannot envy family with lower priorities that are feasible. They show that a stable matching may not exist. What about *initially feasible envy* ? Could we say that a demand that cannot be satisfied *ex-ante* should not be considered in the application ? What does it change if we do that ?

a) Definitions and intuition Let us formally define **initial feasibility**:

Definition 1.12 (Initially feasible application)**.**

We say that a school s is **initially feasible** for family i if before starting the assignment:

- **case without soft quotas**
$$\nexists\, t \ s.t. \ d_i^t > q_s^t$$
- **case with soft-quotas**
$$\nexists\, t, g \ s.t. \ d_i^t > \underline{q_s}^{tg}$$

In plain words, school[47] s is initially feasible if there is no seat/day that family i demands that is not available at school s[48]. Note that since $d_i^t \leq 1$, it implies just that $q_s^t = 0$. In particular, if $\forall s, t, q_s^t > 0$ then the problem is initially feasible.

From there, we define a new criterion of stability that we call **initially feasible envy**:

Definition 1.13 (Initially feasible envy)**.**

We say that applicant i has **initially feasible envy** towards i' if there exists a school (or bucket) s s.t. $s \succ_i \mu_i$ and

1. school s is initially feasible for i and
2. there exists applicant $i' \in \mu_s$ with $\pi_{i'}^s < \pi_i^s$.

If either of the two conditions 1) and 2) does not hold, then there is no feasible envy. In particular, this definition tolerates justified envy of applicants who are not initially feasible. This deviation is *deemed* harmless because these applicants could never have been accepted in this daycare.

To get an insight from what we intend to achieve, we motivate the following result with Example 1.5 in Appendix F.II. This solution was developed to limit vacancies in second rounds of assignment. Indeed, there are often many daycares with empty seats on some days and we saw that parents who could not be seated *ex-ante* were blocking feasible assignments. This example suggests that if we tolerate infeasible envy, then we may (weakly) improve the assignment of all applicants while ensuring no initially feasible envy.

We want to see whether this particular result is valid in the general case. In other words: can we prove that the SOFM with initially feasible demands is weakly better than the SOFM with full preferences for all individuals ? Does this result also apply to weakly envy-free assignments ?

[47] Or the group/seats g at school s
[48] Or at the bucket g at school s

b) Matchings free of initially (weakly) feasible envy improve assignment for everyone Consider a daycare assignment mechanism with multidimensional constraints, its problem $P = (I, S, B, \mathbf{q}, \underline{q}, \theta)$ with student preferences $\succ_i$. Let $\phi := P \to \mu$ be the student optimal fair assignment (SOFA) returned by algorithm ϕ. Let $\bar{\succ}_i$ denotes initially feasible demands. Then we have the following theorem:

Theorem 1. *The assignment μ' obtained with the same mechanism ϕ on* **initially feasible demands Pareto dominates** *μ. We call this equilibrium the* **Student optimal initially feasible fair assignment** *(SOIFFA)*

See proof in Appendix F.II.

Theorem 1 shows that we can assign a (weakly) better outcome to all parents if before the assignment we adjust their preference order and remove demands that cannot be satisfied *ex-ante*. In the same spirit, Kesten (2010) proposes an efficiency-adjusted deferred acceptance mechanism (EADAM), that allows a student to consent to waive a certain priority that has no effect on his or her assignment. Here, policymakers modify parents reported preferences and cause them no harm, but may help many others benefit as a consequence. Although it seems like a harmless *low hanging fruit* to reduce vacancies and satisfy more families, agents may value the idea that all their choices were considered in the assignment and on the other side, policymakers may be reluctant to such interventions. On the other hand, since everyone gets a *weakly* preferred outcome, the adjustment (weakly) improves welfare for all agents. In a setting where the number of choices is not restricted, this adjustment can make participation to the assignment mechanism *safer* in the sense that if a choice is doomed to be rejected it is discarded *ex-ante* and does not negatively affect others. This result is almost ineffective in the first rounds of assignment when there are lots of seats for every day in almost every bucket. In the second rounds, this theorem proves useful.

B) General solutions to the daycare assignment problem

So far, we did not consider student groups and buckets in the definition of the problem but empirical application requires a solution for the problem with multidimensional and diversity constraints. Again, our overall strategy is to transform the problem with diversity constraints in an equivalent and well defined problem that can be solved with CAM or KDA. Before, the "*trick*" was to break diversity requirement into unidimensional unitary buckets. Now, since each school has multidimensional capacities, *i.e.* one for each day, this is not possible anymore. Moreover, we don't want unitary buckets since our goal is precisely to try and combine as many demands as we can over weekdays. In fact, we need a definition of buckets that contains **as many capacities as possible**.

Buckets are defined by policymakers and daycare providers and *should* correspond to empirical constraints (a specific room, equipment, a staff-to-children ratio and so on). They constitute the bundle of capacities across days over which we can combine demands following priorities. **Capacities across buckets cannot be combined**. For our purpose here, we do as before: a bucket b possesses elements $(s_b, \succeq_b, c_b)$ but where now capacities c_b are defined over days so that $c_b(t)$ is the number of seats in bucket b for day t. We let $\tilde{\mu}$ be a matching of students to buckets.

Example 1.3 (Capacities in different buckets)**.** Consider a school s with at least two buckets k and l in B_s and let $c_k = (1,1,0,0,0)$ and $c_l = (0,0,0,1,1)$. Student i asks $d_i(t) = (1,1,0,1,1)$ and both buckets accept student i although not in the same order[49]. In our definition, student i will apply to each bucket and regardless of the definition of envy we choose, i cannot be accommodated in any bucket. However, in school s, there appears to be enough capacity to accommodate i across buckets. In our setting we assume that if capacities cannot be bundled in the same bucket, that is because a child cannot be some days in one bucket, and the other days in another, even if student i is eligible to both buckets.

[49] Otherwise, they would have been aggregated together.

a) Stability definitions in the problem with multidimensional constraints and soft quotas Without buckets, we introduce the **strong envy-freeness** definition of Kamada and Kojima (2023) and the **weak envy-freeness** notion of Delacrétaz, Kominers, and Teytelboym (2023). The main difference lies in the definition of feasibility. where the latter allows *weak accommodation* and the former does not. Both rely on priorities. Our Theorem 1 adjust the set of preferences to limit externalities but rely on the same definition otherwise. All the following results can use initially infeasible demands. Let us give a formal definition of envy-freeness with diversity quotas:

Definition 1.14 (Strong justified envy with diversity quotas)**.** Given a matching $\tilde{\mu}$ of students to buckets for the problem, applicant i in I from group $g_i = g$ with assignment $\tilde{\mu}_i = b$ has *strong justified envy with diversity quotas* over family $j \neq i$ of group $g_j = g'$ with assignment $\tilde{\mu}_j = b'$ if and only if

1. $s_{b'} \succ_i s_b$: *Student i prefers the school of the bucket of j* **and**
2. $\tilde{\pi}_{ib'} \geq 0$: *Student i is eligible for seats in b'* **and**
3. $\tilde{\pi}_{ib'} > \tilde{\pi}_{jb'}$: i has higher priority score in bucket b' than student j and $\tilde{b'}$ cannot weakly accommodate j alongside with all families with higher priority than j at b' that weakly prefer $s_{b'}$ to the school of their assigned bucket.

The definition is simply a version of weak envy-freeness in a matching of students to buckets where only the preferences over schools matter to define whether a student has a strong justified envy.

b) Assignment mechanisms for problems with multidimensional constraints and diversity quotas Policymakers have important choices to make daycare assignment marketplaces functional and we want to emphasise what it means in terms of fairness, optimality and strategy proofness. Our main message can be wrapped up in Definition 1.15 and Theorem 2:

Definition 1.15 (Daycare assignment marketplace (DAM))**.**

A **daycare assignment marketplace** (DAM) is built upon three elements policymakers must **define**:

1. A **version of the problem** : considering demands with specific days or not ;
2. A **partition of capacities** into buckets ;
3. A **definition of *fairness*** : whether they want an envy-free or weakly envy-free assignment, on initially feasible demands or all demands.

The definition of a DAM fixes the problem $P := (I, S, B, \mathbf{q}, \underline{\mathbf{q}}, \theta)$ which follows definition 1.5. Using definitions 1.3 and 1.4 to create new priorities and preferences with buckets, we map elements of P in a simpler problem. The definition of fairness also imply the target properties on assignment μ and therefore, the appropriate mechanism ϕ.

Hence, once policymakers fix a DAM, we propose the following theorem:

Theorem 2. *For a given daycare assignment marketplace (DAM), there exists a unique assignment that is the most preferred by all families, respects priorities, multidimensional and diversity constraints. We call this particular assignment the* **Student optimal fair assignment** *(SOFA) which can be found by one of the three algorithms.*

Proof. The proof directly follows from the definition of the problem with diversity quotas and the properties of the mechanisms used. In particular, the careful definition of buckets, priorities and preference order over buckets in each model maps the original problem into one that fits the definition of the input for each algorithm. Their properties presented earlier directly apply to the problem with diversity quotas and our definition of envy-freeness with diversity quotas.

More precisely:

1. The SPDA algorithm returns the only non-wasteful strategy-proof SOFA for DAM with diversity constraints.

2. The CAM algorithm returns the only SOFA for DAM with multidimensional constraints, diversity constraints and fairness defined by envy-freeness.
3. The KDA algorithm returns the SOFA for the DAM with multidimensional constraints, diversity constraints, and fairness defined by weak-envy freeness.
4. From an assignment with CAM or KDA on a DAM with multidimensional constraints and diversity constraints, the same mechanism on initially feasible demands Pareto dominates the assignment with all demands.

QED

Definition 1.15 emphasises how the design of an assignment depends on a set of political decisions. The definition of the problem has consequences on what is feasible, on achievable properties and so on. When policymakers define a DAM, we provide clear solutions with traceable assignments. All models return student-optimal matchings. There are no better assignment that at least one family prefers that respects the **chosen** definition of *fairness*. This is true with our definition of buckets as we cannot feasibly accommodate students without breaking the definition of buckets and the associated group ranks.

When one defines a DAM, many details will have consequences. In particular, the partition into bucket is a very sensitive task. An envy-free DAM and a weak-envy free DAM will not yield the same outcome, even if the rest of the problem is identical.

c) Extension: allowing mobility of already assigned families In practice, some parents assigned in previous committees may want to participate in the assignment mechanism again in order to obtain another daycare. Some settings allow such mobility although they usually struggle to define a fair way of doing so.

The use of the assignment with diversity constraints makes this operation easy[50]. One simply requires these parents to submit their preferences and municipalities to set their level of priority in the childcare they demand and flag these files in the data. Then, we proceed as follow:

1) Add the daycare they come from as their last school ranked
2) Add the capacity they would use in this daycare to quotas and buckets.
3) Set these student's priority over this bucket so that they (and only they) have the highest priority if they end-up unassigned to preferred school.
4) Run the algorithm corresponding to the DAM.

Combe, Tercieux, and Terrier (2022) discuss in details the properties of a similar assignment mechanism to assign teachers in middle and high schools in France. In practice, we only encountered one municipality that allowed such transfers.

d) Main limit: multiple demands While our models are designed to accommodate individual applicants (students), the reality often involves families submitting multiple demands simultaneously. Twins are easily accommodated as they usually are of the same type and receive priority bonuses. To foster joint assignment, we pick the highest lottery realisation within the family, and assign them the highest (+ $1/10^6$). That way, they always follow each other in the demand sets and separation only occurs when they end up last in the bucket. Assignment committees usually manage to *push the walls* and accommodate them. Otherwise, they unmatch them and they go again in the next round. The problem is harder when families have children of different ages and submit several choices. Handling this situation becomes challenging due to variations in competition across different daycares and the unique constraints each bucket imposes. This practical constraint is not fully considered in our model, and further research is required to address it effectively. In practice, these challenges have significant consequences, as

[50] With initial endowments and priorities, it is well known that there is no matching that is both individually rational and stable in the sense that no agent has justified envy. To ensure the compatibility between the two notions of IR and stability, the concept of stability has been relaxed to exclude blocking pairs caused by a daycare slot that is assigned to its initial owner. A condition known as μ_0 $stability$. See Combe and Schlegel (2021)

committees must sometimes deviate from theoretically optimal fair assignments. Even minor deviations can result in substantial justified envy, making it difficult to individually justify the assignments. We discuss these practical implications in the following section.

In subsection IV.5, we illustrate the tension between *fairness* and *wastefulness* by verifying the student-optimal and envy-freeness properties of our assignments and compare outcomes of different algorithms.

IV.5 Delivering the SOFA: empirical illustration

We focus on the last year of assignment of VRA, the first case study presented in sub-section III.1 in section III. VRA defines the DAM with demands over weekdays, groups by age (rounded in months) on September the 1st. In 2020, they chose a definition of *fairness* based on envy-freeness and from 2021 onward, they opted for assignment on initially feasible demands. We coded all solutions of Theorem 2 in Matlab and used them to assign daycare seats instead of assignment committees. The main testable properties of our assignment are **envy-freeness** on **initially feasible demands** and **optimality** accounting for **diversity constraints**.

One way to show these properties at work is to rebuild the demand set of each bucket, retrieve the cut-off of the last accepted file and observe each file's status[51]. The properties have implications for the content of this table. In particular, families with rankings above the threshold of the last admission can only be allocated to this section or to a daycare centre higher up on their preference list. Families ranked below this threshold cannot be accommodated in this bucket to prevent justified envy. If the daycare facility has flexible age groups, they might find placement in another bucket of the same school eventually. Otherwise, they will either be allocated to a different daycare centre (the highest one on their preference list where they are eligible) or placed on the waiting list. In practice, this step serves both as a verification and means to build justifications for each assignment[52].

A) Student optimal fair matching and transparency

For every buckets $b \in B$, let $\underline{c}_s$ be the **cut-off** of bucket b which is an integer in $\{0, \dots, N\}$ where $N = |I|$ is the number of applicants. Once ties are broken and the scores $\widetilde{pi}_{is}$ at bucket s are obtained, order the applicants by decreasing order of score and let $\tilde{r}_{ib}$ be the **rank of applicant** i in this ordering.

Given a collection of cut-offs $c := (c_b)_B$, let:

- $A_i(c) := \{b \in B : \tilde{r}_{ib} \leq c_b\}$ be the set of **accessible buckets** for applicant i.
- $D_b(b) := \{i \in I : b \in A_i(c)$ and $b \tilde{\succ}_i b'$ for $b' \in A_i(c) \backslash \{b\}\}$ be the **demand set** of bucket b at the cut-offs c.

In plain words, accessible buckets are those where the rank of an applicant i in that bucket is lower than the cut-off ; the demand set is composed of all applicants in their most favourite accessible bucket.

Figure 1.5 presents a proportional stacked bar chart illustrating demand sets rank relative to the last admitted file in each bucket. This chart verifies and emphasises the key features of our models. For policymakers, these properties can be summarised in Fact 1.

Fact 1. *Each registered family was assigned to their most preferred daycare centre from a set of options where they met the assignment criteria. This allocation was determined by considering initially feasible demands, which were assigned by preference order following priorities within age groups and subject to capacity constraints over different days.*

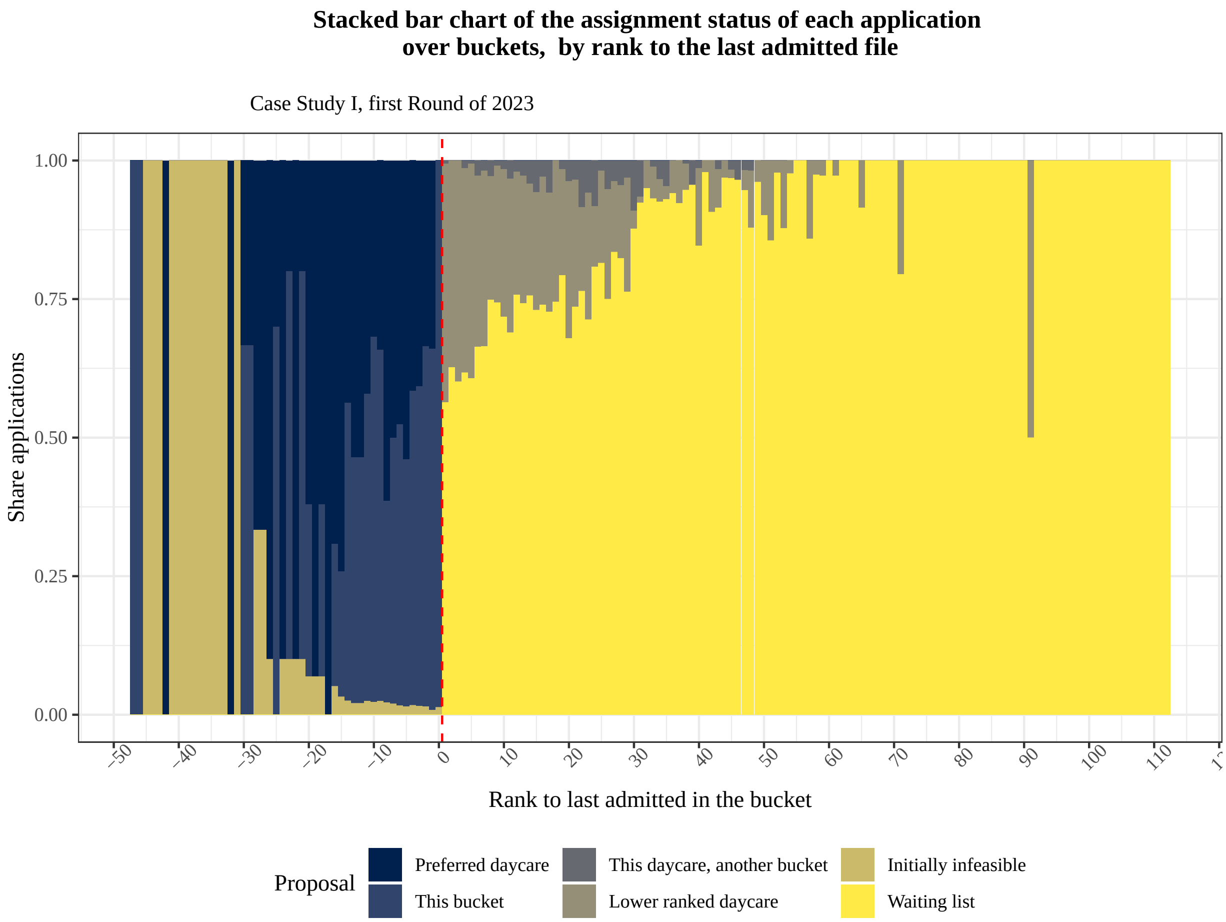
[51] In practice, we start by selecting the 1267 records from the assignment file of this round and pivot it over each reported choice. This file is made of 5005 reported choices. Next, we match each choice to the corresponding buckets. Using age bounds and flexibility criteria between age groups, we determine which choices are eligible for each section and exclude those that are not. Once this step is completed, we have a database of 3740 choices over buckets. From there, we can reconstruct the actual priorities by ranking families within each bucket based on the priority age group, then the score, and finally the lottery number. We retain the order of files in each bucket and keep the ranking of the last admitted in each section.

[52] Each, year we provide an Excel file with one sheet per daycare. In each sheet, we provide the demand set over each bucket and all relevant information to justify assignment. In particular, we show the decreasing number of available seats for each day, and the assignment of every file.

While initially infeasible demands are less common in first-round assignments, they still exist. Some age categories with older children have certain days with zero available capacities and some initially infeasible preferences were removed before the assignment. Had we retained them, 6 fewer children would have been assigned. In the second round, where zero capacities are more prevalent, the welfare improvement of SOIFFA over SOFA becomes much more salient. During that year, our assignments increased from 100 to 115 when considering initially feasible demands, which represents a 15% rise in admissions.

This last finding underscores the significance of a precise definition of *fairness* in the daycare assignment marketplace. This is especially crucial if policymakers aim to minimise inefficiency while upholding the principles of due process.

Figure 1.5: Student-optimal initially feasible fair matching with diversity constraints and initially feasible demands: an illustration



Sources: ISAJE, Case Study I − 2023, First round only.
Notes: This figure is based on initially feasible family preferences. Using age group and score definitions, we calculate the rank of each applicant in each bucket where they applies.
These calculations allow us to verify the properties of the algorithm.
Negative values are above the last admitted applicant's in this bucket.
Among higher ranked families, applications are either in a preferred daycare or in this bucket.
Below, families are either assigned to another daycare, to another bucket of the same daycare when buckets are flexible, or placed on the waiting list.

B) Assignments with alternative procedures and final assignments

The CAM algorithm is the most rigorous with priorities and tolerates no weak accommodation. Allowing deviations for initially infeasible demands improves assignments for everyone (SOIFFA) and can theoretically seat more families. However, the assignment can be wasteful. We illustrate the leftovers by days across daycares after the assignment we just presented in Figure D.25 in the Appendix. There are many days lefts but there are few within buckets and mostly on Wednesdays, the least demanded day. Another interesting question is to compare the assignment with *weak envy freeness* with KDA.

Assessing the cost of envy-freeness Table 1.1 compares the SOIFFA with alternative assignments[53] and the final offer.

The SOFM with all demands is the most restrictive regarding priorities and tolerates no justified envy. It yields 402 assignments and 865 in the waiting list. Removing initially infeasible demands adds 6 families, 2 more first choices, 3 third choices, and one fourth choice. In this market, allowing weak accommodations on all demands gives the same number of seated families as removing initially infeasible demands. However, KDA serves more second choices and weakly dominates the SOIFFA.

Table 1.1: Comparison of assignments with different procedures and the final assignment

Results	SOFM All demands	SOFM Initially feasible	KDA all demands	Brut force	Final assignment
N 1st choice	238	240	240	264	**258**
N 2nd choice	85	85	87	98	**96**
N 3rd choice	46	49	48	63	**52**
N 4+ choice	33	34	33	42	**33**
Waiting list	865	859	859	800	**825**
Total offer	402	408	408	467	**439**
Total	1,267	1,267	1,267	1,267	**1,264**

Sources: ISAJE, Case Study I - 2023, first round only.
Notes: Columns 2, 4, 5 are simulations based on the same files.
Column 1 is the SOFA obtained using the CAM algorithm on all demands.
Column 2 is the SOIFFA delivered to the early childhood department.
Column 3 is the SOFA obtained using the KDA algorithm on all demands.
Brut force simulations assign as many feasible demands as possible without following priorities. They give an upper bound of the total number of children that could be assigned without priorities.
Final assignment indicates the proposals made to the families.

[53] SOFM uses the CAM algorithm, the SOFA returned by KDA uses weak-envy freeness to define fairness.

Respecting priorities induces some rather large vacancies. Filling as many vacancies as possible by brute force[54], one could have assigned 467 children. Fairness *cost* 59 vacancies that could have been filled. However, this assignment has absolutely no property apart from leaving less (but not necessarily zero) empty seats.

However, the assignment committee observes these vacancies, and in the final assignment, they assign 31 more files than the SOIFFA. Some of these changes are quite inevitable, particularly because in the current use, we are not able to guarantee that multiple demands by the same parents are assigned the same daycare. In this example, 29 of the modified applications correspond to multiple demands. Since the procedure still assigns a maximum number of applications, moving around thirty of them forces the reassignment of others. In the end, the early childhood committee modified the allocation of 94 applications.

Adjustments and deviations from envy-free assignments Table D.9 in the Appendix compares the deviation from the SOIFFA in the final assignment. The committee added 47 files from the waiting list to assign them a place and improved the assignment of 17 files. However, to make these adjustments, they had to revoke assignments for 15 families and downgrade 15 others to less-preferred childcare centres in their preference order. These adjustments inevitably introduce justified envy, and not just from those who were moved, but anyone with lower priority in the demand set. Therefore, the assignment no longer respects the priorities established by the territory and we cannot justify the allocation of moved files based on the rules of the procedure. These adjustments have implications for transparency. However, we believe that this issue can be resolved by defining conditions under which the committee can deviate from priorities. One possible solution would be to establish a voted, public, binding ethical charter. This would allow the allocation committee to make deviations from priorities for specific and legitimate reasons, such as “not separating twins”, while restoring transparency to the decisions. The solutions we propose remain tools supposed to support decision-making. They efficiently handle large demands while eliminating technical challenges related to priority compliance, age groups, and capacity constraints. However, access to childcare may require more tailored decision-making in some cases.

V Market design in the field: *Who gets what and why ?*

In what follows, we analyse the functioning of our assignment procedures in the field, with results from the two case studies introduced in section III. We use our unique position of *plumber, engineer and scientists* (Duflo 2017) to analyse the consequences of their main features. First, we focus on the Case Study I to analyse the role of priorities using simulated assignments with different priority rules. Then, we detail the consequences of time since registration. Finally, we use the second case study to show the effects of buckets.

As discussed in section II, The evaluation of a market design require normative criteria. It hinges not only on its consequences but also on whether these consequences are desirable, undesirable, intentional or not. In the face of any consequence, we may be called upon to make normative judgements. Are we morally permitted to bring it about? Are we morally obliged to bring it about? As noted by Li (2017a), “*informed neutrality requires us to engage with reasonable ethical theories, not to resolve which one is correct, but to formalize the features that they regard as morally relevant*”. In this research, our *moral compass* is the Matthew effect and our analysis aim at understanding how DAM foster or mitigate it.

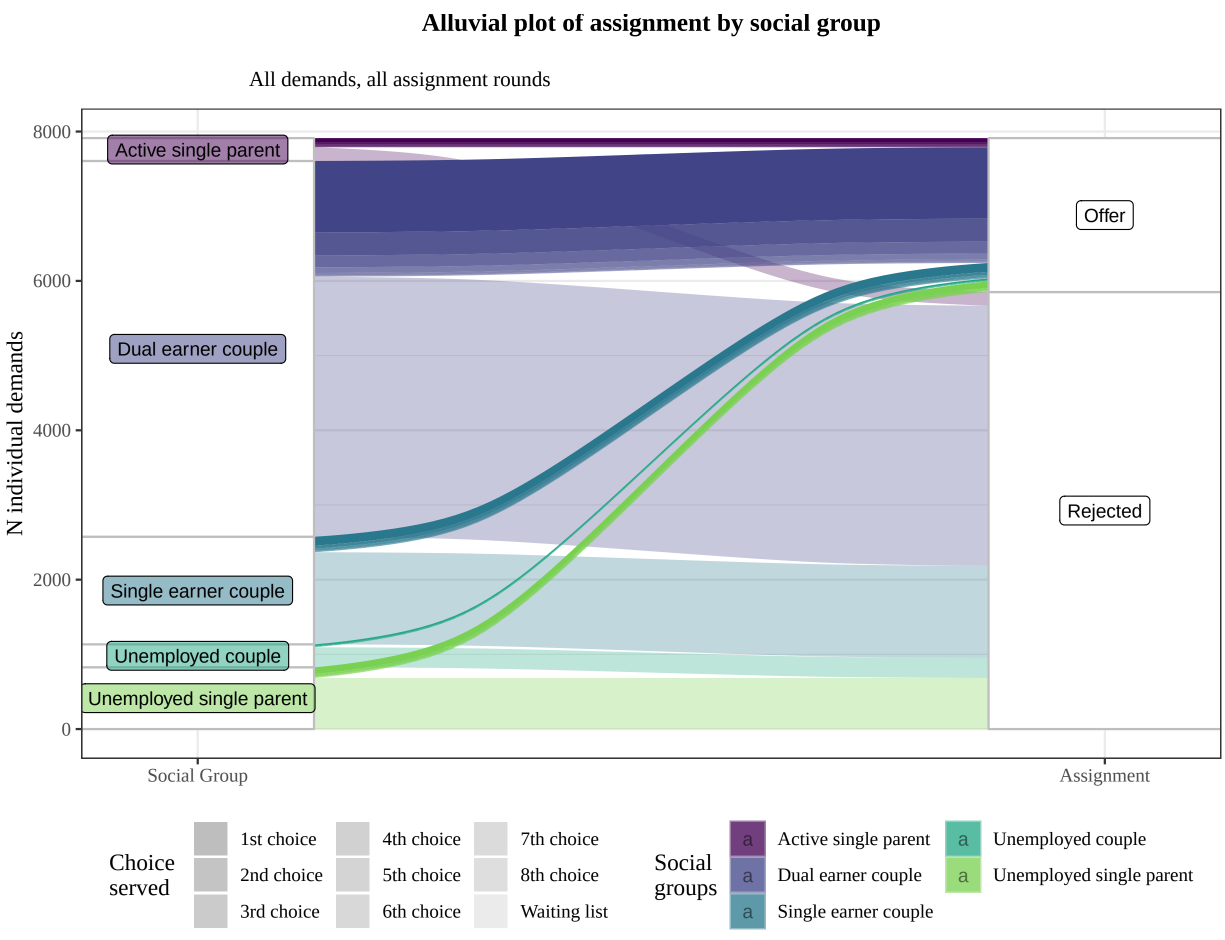
[54] The brute force algorithm simply fills as many children as possible proceeding as followed. 1) Sort Buckets, 2) **For** each bucket **do** : i) Order families. ii) Match if seats available. iii) Move to the next family. 3) Move to the next bucket. 4) Repeat Until Done. **EndFor**

V.1 Case study I: Who gets what ? General overview

Main results of the assignments Each year, we processed an average of 1144 applications in the first rounds and 834 in the second rounds (See table D.7 in the Appendix). In this first year, we allocated seats to 400 families, which accounted for approximately 40% of the total demand. About two-thirds of seated families received their first choice. The outbreak of the pandemic resulted in minimal new registrations for the second round, during which we successfully assigned an additional 111 children. Over the years, the competition intensified, demand surged in 2022 and despite more available capacities, the proportion of accepted families got lower and lower. On average, 36% of demands were assigned a seat in first rounds.

In terms of social composition, 64% of all files come from dual earners, 18% single-earner couples and 4% couples where neither work (See Table C.5 in Appendix A). In this market, active single parents represent 4 % of all demands, and single parents with no job 10 %. Figure 1.6 summarises the distribution of assignments across all rounds and years by social group. Overall, double earner couples get 75% of all seats while representing 64% of all files.

Figure 1.6: Flows of assignment by social group



Sources: ISAJE, Case Study I − 2020 : 2023.
Note:Number of assigned and rejected files in every assignment over the four years by social group.

Table D.8 present the average characteristics of seated and waitlisted families, showing significant differences in virtually all included measures. Families who submit only one wish, on average, have a 6-point lower probability of being admitted. On the other hand, those who submit at least three wishes are 10 points more likely to receive an offer than those who rank fewer daycares. Admitted families tend to enroll much earlier, with a significant number doing so before delivery. For instance, admitted children are, on average, one month older than those on the waiting list. Moreover, children who are admitted in September have been registered for an average of 13 months, which is

3.5 months longer than those on the waiting list, on average. It is also worth noting that the average child age at the time of registration is 2.5 months younger among assigned families than among waitlisted ones. Additionally, 65% of admitted children were enrolled before birth, compared to 47% of rejected children.

Interpretations in the context of the pandemic Among other things, the economic uncertainty brought about by the Covid-19 pandemic had large impacts on family dynamics[55]. The INSEE (2021b) collection show that the first lock-down led to a significant decline in births. This is especially true for places heavily affected by the health aspects of the pandemic but less so economically, such as VRA. Births between December and February in this region decreased by 13.37% compared to the three previous years' average for the same period. There was a slight increase in births, by only 2.55%, from March to May[56].

In our report for the local administration, we describe in details an unexpected effect of the pandemic that changed the distribution of registrations and birth dates, causing a reduction in differences between social groups in the 2021 assignment[57]. Figure C.17 in the Appendix plots the cumulative distribution to show stochastic dominance. In 2021, with these enrollment lags, the empirical cumulative distribution of priorities for single parents and dual-working couples overlaps on the full support of priorities and 1/4 of family without work have higher priorities than half of dual earner couples. Said differently, the scores no longer allow to sort groups along policymakers preferences. Facing more complains from working families, they revised it in 2022, giving a massive priority boost to dual-working couples and active single-parent families.

V.2 The 2022 priority reforms and its consequences

The 2022 reform simply consists in adding 5 more points to dual earner couples and single parents working leaving everything else the same. The former went from 3 to 8 and the latter from 5 to 9, while cumulating priorities of other criteria. In 2022, the mean priority score for other group is 8.3. Both groups have thus been given higher priorities than the average priorities among other parents. However, parents were not informed of this change and could not anticipate its effects. In spring 2022, the new weights were simply implemented in our algorithms and we were not involved or informed before we received the data. The boost is so massive that **only 30 % of single parents without a job have a higher priority than the 10 % dual earners with the lowest priorities**. However, the 2022 demand is much larger than the previous years, with more demands from couples with one parent working and single parents, more often for part times for both groups (See Figure C.20 in the Appendix). Most of them being waitlisted, a large share remained registered. Assuming that the *ranking* of weights across social groups can be interpreted as policymakers' preferences, Table D.8 (in the Appendix) shows that the differences between assigned and waitlisted families satisfy policymakers' preferences on average. Yet, they were disappointed again by the distribution of assignments in 2023 as the share of assigned working parents shrank, despite a large increase in their priorities in 2022. In what follows, we analyse how and why by answering a series of *what if* questions on the role of priorities.

[55] For instance, Barbuscia et al. (2023) show large increases in separation rates, especially among young couples, in the six months following the initial lock-down. Individuals who were unemployed or had lower incomes before the pandemic are more likely to separate immediately after the lock-down. Degraded financial situation for men is associated with a higher risk of separation throughout this period. Lu et al. (2022) reveal a continuous rise in anxiety and depression symptoms during lock-downs in France. A small group of people experienced a significant and long-lasting deterioration in their mental and physical health. Breton et al. (2021) document major structural demographic changes, including a decrease in births, marriages, immigration, and an increase in deaths.

[56] Among families registered for daycare in our data, the drop and rebound are almost of the same magnitude.

[57] In brief, registration were not possible during the first lock-down and no national political orientation regarding the reopening of daycare centres were made before late August 2020. There are far fewer families registered between March and August compared to the other years and therefore less variation in priorities. Furthermore, the pandemic affected the timing of births. 9 months after the first lock-down, there are far fewer children born compared to other months but many pregnancies were simply postponed and there is a surge of children born in spring 2021.

A) What if there were no or other reforms ? Simulated counterfactual analysis

In this setting, if we had been given different priority parameters, we would have treated the assignment problem the same way with these alternative priorities. Policymakers could have asked the effects of alternative priorities before voting on the 2022 weights[58]. In the Neyman-Rubin framework, causal analysis is often seen as a *missing data* problem where only one potential outcome can be observed (Imbens 2020). Other approaches of causality use theory to either draw causal graphs or structural econometric models that impose sufficient condition for identification (J. J. Heckman and Pinto 2022b). Here, the economic model is the actual model. We can actually **know** the missing individual counterfactual outcomes by simply running our algorithm with alternative priority rules, holding everything else constant. In particular, we keep the same lottery realisation so that ranks only vary with the change we test. In particular, we ask:

1) **What if the score was not changed ?** We simulate a "Control" assignment with the former weights with the 2022 data;
2) **What if there were no priorities at all, only age groups ?** We simulate a "Random" assignment using only the random tie breaker to sort applicants for assignments in 2022 and 2023;
3) **What if priorities stopped considering time since registration ?** We simulate a "Social" assignment using all priority criteria but time since registration for assignments in 2022 and 2023;

Our strategy ends-up very close to the model of Kasy and Abebe (2021) on the causal impact of algorithms on inequalities. They propose a 4-step algorithm audit using distributional decomposition: i) Normative choices: what are the relevant outcomes, measures of welfare and/or inequalities, and so on ii) Calculation of influence function: using a machine learning algorithm, iii) Causal effect estimations and iv) Counterfactual assignment probabilities.

They require influence function because their framework is based on predictive algorithms. In our setting, we do not need these estimations and can directly predict individual counterfactuals instead. As normative choice, we analyse how the assignment probabilities conditional on social groups evolve and later, we analyse segregation.

We take a simple potential outcomes notation and let $Y_i = Y_i(T)$ be the observed assignment with the reform. Let $Y_i(c) \quad c \in \{C,R,S\}$ be the counterfactual assignment in the Control, Random and Social scenario. We note $D_i = \mathbb{1}(d = c)$ and X denote social groups. For each social group x We want to estimate :

$$\mathbb{E}[Y_i(c) - Y_i(T)|X = x]$$

$I.e.$ the expected difference in assignment probability between the observed and counterfactual situation. For simplicity[59], we define Y as a dummy for being assigned. To estimate these, we simply build a stacked database for each values of D and actual and counterfactual assignments and analyse the differences like a match-pair design. We have two observations for each pair of individuals i, we simply estimate the following equation separately by social groups using OLS and cluster robust standard errors at the pair level:

$$Y_{idx} = \alpha_i + \beta_x \mathbb{1}(d = c) + \varepsilon_{idx} \quad \forall i \text{ s.t. } X_i = x \tag{1.1}$$

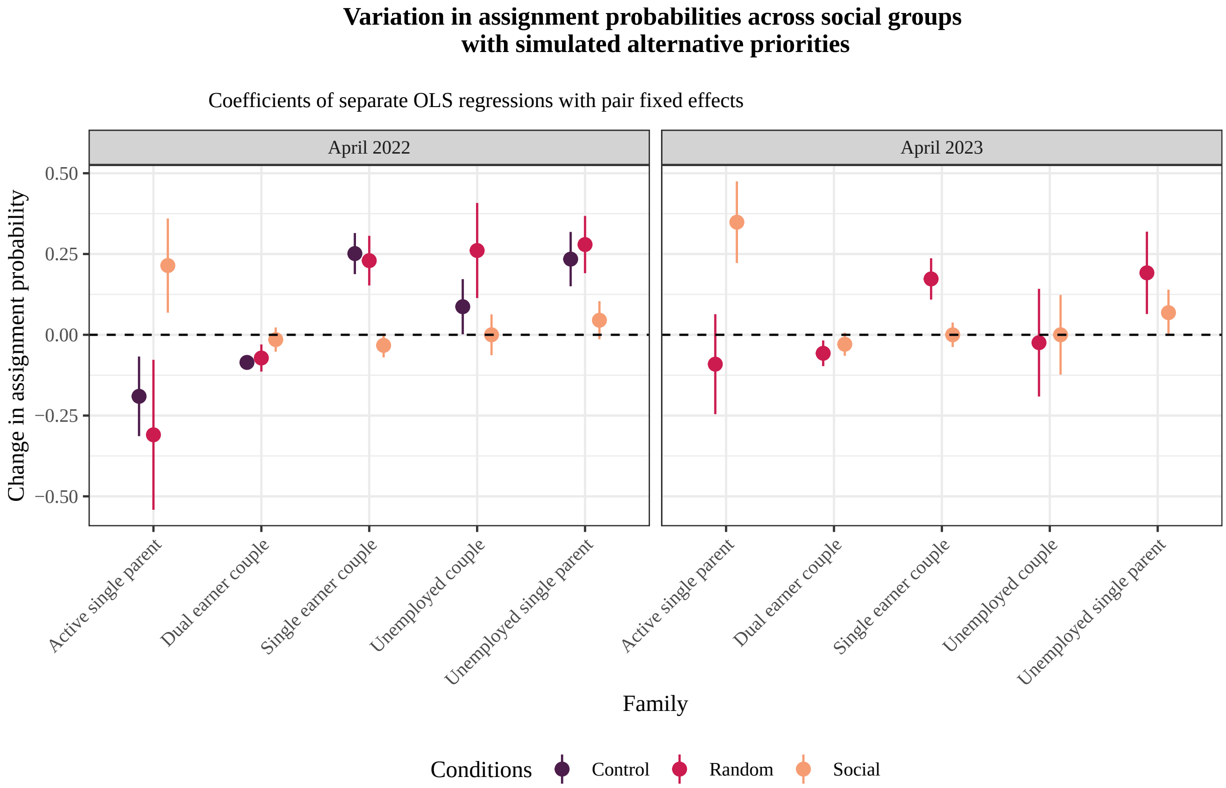
Where β_x retrieves a weighted average of within individual difference in assignment from observed to simulated. The cluster robust variance accounts for the correlation within individuals. Since nothing else changes between the two conditions but the priority parameter we manipulate, we interpret these estimates as the **effects** of the priority we test. For instance, when we compare the average difference between actual and social scenario for active single parents, we measure the average effect of *suddenly* removing the early registration premium on the average probability of a single parent to be assigned. Confidence intervals are clustered by pair of individual and used to gauge variability due to group size using large sample approximation for simplicity[60]. We present the results of these estimations for 2022 in Figure 1.7.

[58] In fact, they did ask for such simulations after the first year. At that time, the debate was about setting priorities for parents asking daycare in the same city they live. The question was about territorial segregation and our analysis showed that such change would i) drastically reduce assignment probabilities of parents living in cities without a daycare and 2) strongly favour residents of the two largest cities. This reform project was dropped after we presented our evidence.

[59] We could test more precise hypotheses like better or worse assignments, same assignment and so on. While these are interesting for policymakers, we don't think they would bring much in this research paper.

[60] Note that a reform where parents would know that time since registration does not matter would affect the composition of the demand. The external validity of these estimates strongly depends on the self-censoring and strategic timing this criteria generate among parents who consider daycare. An important question that we will be able to adress matching these data with those of the National Family allowance fund.

Figure 1.7: Effects of priority criteria on assignment probability by social groups



Sources: ISAJE, Case Study I − 2022 : 2023 − first rounds only.
Notes: We simulated assignments using alternative definition of priority scores while all other parameters are held constant.
We stacked simulations in a database and present the coefficients of separate regressions by social groups of a dummy for assignment over scenario dummies and individual fixed effects. Confidence intervals are clustered at the individual level.
'Actual' assignment is the reference, 'Control' is the scenario without the reform i.e. the 2021 weights,
'Random' is the scenario without priorities i.e. students are sorted by decreasing lottery realisation with age buckets,
and 'social' use the priorities without weights for time since registration. In 2022, it uses reformed weights.
We use cluster robust standards errors at the individual level to build 95% confidence intervals.

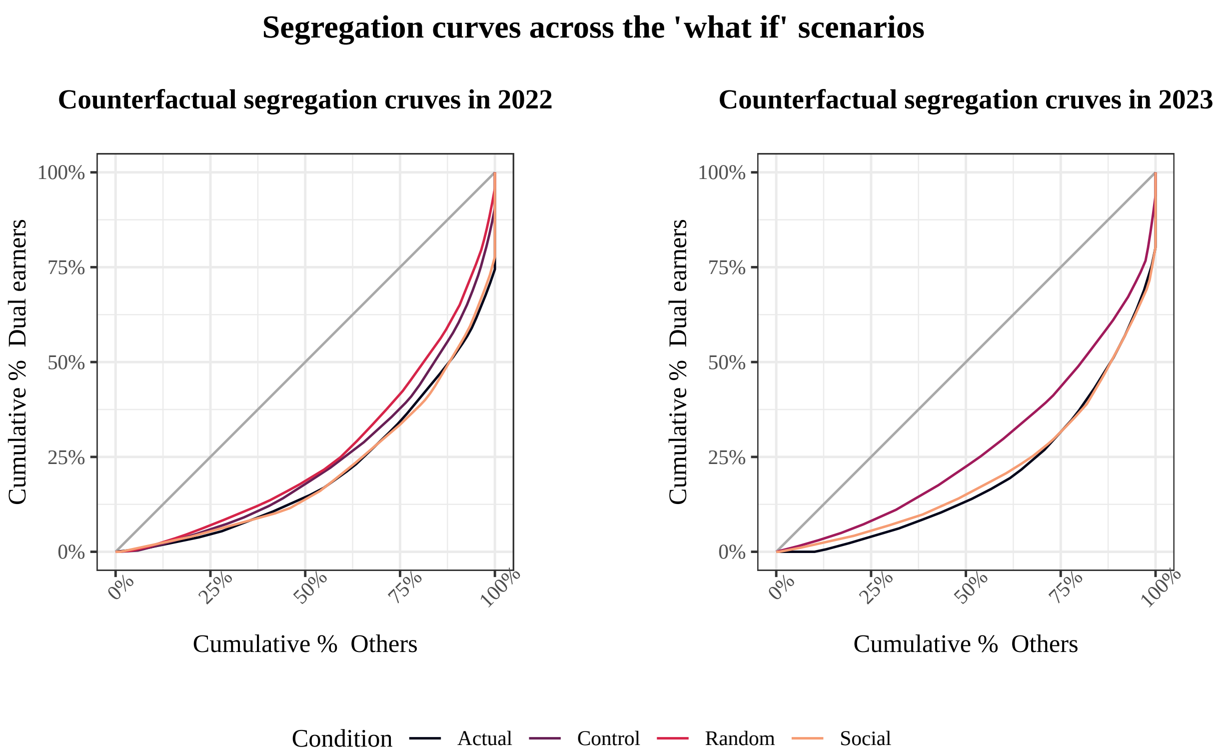
First, the comparison of assignments with "Control" shows that the reform increased the share of dual earner couples and active single parents, but at the expense of all other groups. Second, the results for the simulation without priorities ("Random") are very close to the control ones, except unemployed couples. Third, removing time since registration ("Social") has almost no aggregated effects across social groups but for single parents. Working single parents are most heavily penalised by time since registration while the effect is smaller for single parents without a job. Both results also hold true on the 2023 data.

Such simulations provide valuable insights for policymakers and can help build better weighting schemes. Our goal is not to criticise the politics of favouring working parents but our tools can help answer whether it is effective, who benefits, who loses, and by how much. The reform is highly effective the year of its implementation. However, time since registration priorities do not affect assignment probabilities across groups except reducing that of single parents. In 2023, time since registration reduces the effect of the reform on the group with the highest social weight.

B) The effects of priorities on segregation within daycares

The conditional assignment probabilities and effects of the alternative priorities are estimates over all demands but they may not be the relevant measure. Perhaps policymakers' are more interested in the distribution of social groups among those accepted and within daycares. Segregation measures describe such differences in the distribution of social groups across daycares, and there are many indicators capturing different aspects of the same phenomenon (Mora and Ruiz-Castillo 2011). First, we simply uses segregation curves[61] proposed by O. D. Duncan and Duncan (1955) in Figure 1.8. The left panel for 2022 shows that the reform increased segregation. That was the intended effect. We also see that the segregation curve of assignments without time since registration in the priority overlaps almost perfectly. On average, the priority scores would yield the same level of segregation between dual earner couples and other groups without time since registration. Without the reform, the segregation curve would have been very close to that of assignment without priorities. The right panel shows that in 2023, the two segregation curves with priorities are way below that of assignment without priorities showing that priorities do increase segregation. Without time since registration, the assignment would be slightly less segregated but otherwise keep the higher advantage for dual earner couples.

Figure 1.8: Effects of priority criteria on segregation



Sources: ISAJE, Case Study I − 2022 : 2023 − first rounds only.
Data collapsed by scenario, daycare assigned and social groups.
We compute the proportion of dual earner couples and other social groups in each daycare in each assignment.
We order daycares by increasing value and compute the cumulative share of each group.
Segregation curves defined by Duncan and Duncan (1955).

Finally, we look at segregation within daycares using the new visualisation method of *"segplots"* proposed by Elbers and Gruijters (2024). Intuitively, it sorts daycares on the x axis based on their level of segregation measured by a local version of the Theil Information Index (H index). Binwidth are proportional to the number of assigned families and the proportions represent the share of each social group in each daycare. In the bottom panel, we also

[61] A segregation curve represents the cumulative fraction of dual earner couples (on the y axis) and the cumulative fraction of other type of households (on the x axis) when groups are ordered from those with low values of the relative share of each group to those with high values of the ratio (Hutchens 1991). Like Lorenz curve, perfect integration implies a segregation curve on the 45° line. A segregation curve 'dominates' another if it lies at no point below and at some point above the other. The distribution associated with the dominant curve is ranked as more equal.

report each daycare h index and the overall H index. We use this new method to represent segregation over social groups across daycares for the actual and 'what if' scenarii in 2022 (Figure D.26), and the actual segregation over the four years (Figure D.27). Results are presented in the Appendix. The comparison of segregation across 'what-if' assignment and the actual assignment shows that i) segregation is highest with the reform and many daycare only get children from dual earners. Removing time since registration allows more active single parents to be assigned but the segregation remains high. Conversely, segregation without reforms and without priorities both give much more diversity in most daycare and the H index is 10 points lower than with the reform. The new priority weights for dual earner couples increased segregation in 2022 compared with simulated counterfactuals. We can see that segregation remains high in 2023 and was much lower in 2020 and 2021 (Figure D.27).

Interpretations The prioritisation of dual earner couples and active single parents in 2022 led to an increase in their assignment probabilities and a higher degree of segregation within daycares. Policymakers were content with this distribution, and the 2023 assignments have maintained a similar level of segregation across daycares. Our research reveals that when time since registration is removed from the equation, the assignment probabilities remain consistent for all groups, except single parents, who are disproportionately crowd-out because of more recent registrations. On the other hand, our findings suggest that time since registration primarily reallocates assignments within social groups. As dual earner couples make up two-thirds of all demands, their sorting has a significant impact on the placement of other groups. Given this, the timing of registration is a crucial factor for these couples, as we will illustrate in the following subsection.

V.3 Parents' needs and strategic timing of registration

In this section, we study the strategic use of time since registration across social groups and their consequences on assignment probabilities across birth months. This criteria is very often used in the childcare market. It is simple, easy to understand and to explain, and *queuing* usually is the way to go through administrative process. It also shares the responsibility of priorities with parents. Without time since registration, priority scores are coarse and policy-makers let parents compete by signalling anticipation and willingness to wait. In a way, it reduces policymakers accountability because *late* parents can only blame themselves.

Our main argument is that time since registration is a highly unfair criteria mostly penalising vulnerable families, especially single parents through two main mechanisms. First, strategic parents create negative externalities on other families ; second, groups with lowest social priorities are typically those who cannot anticipate their needs and are penalised twice.

Strategic registration: Bunching registrations at pregnancy declaration Dual earner couples often request childcare earlier and mostly compete among themselves through early registration. They represent the majority of the demand and only active single parents have higher social weights. Parents, while aware of the four-priority criteria[62], have no precise knowledge about *how* these criteria are used. What they do know is that time since registration is important and that they can register from the declaration of pregnancy[63]. Note that this is a local feature. In many other marketplaces, registrations are typically accepted after the 7th month, or sometimes after birth. Regardless, when one uses time since registration for priorities, a lower bound for anticipation always exists and latecomers are penalised by design. Knowing that, parents face two significant challenges:

1. **Strongly anticipating their needs**: Parents are compelled to make decisions about their childcare arrangements at a remarkably early stage or risk being crowd-out by other strategic parents.

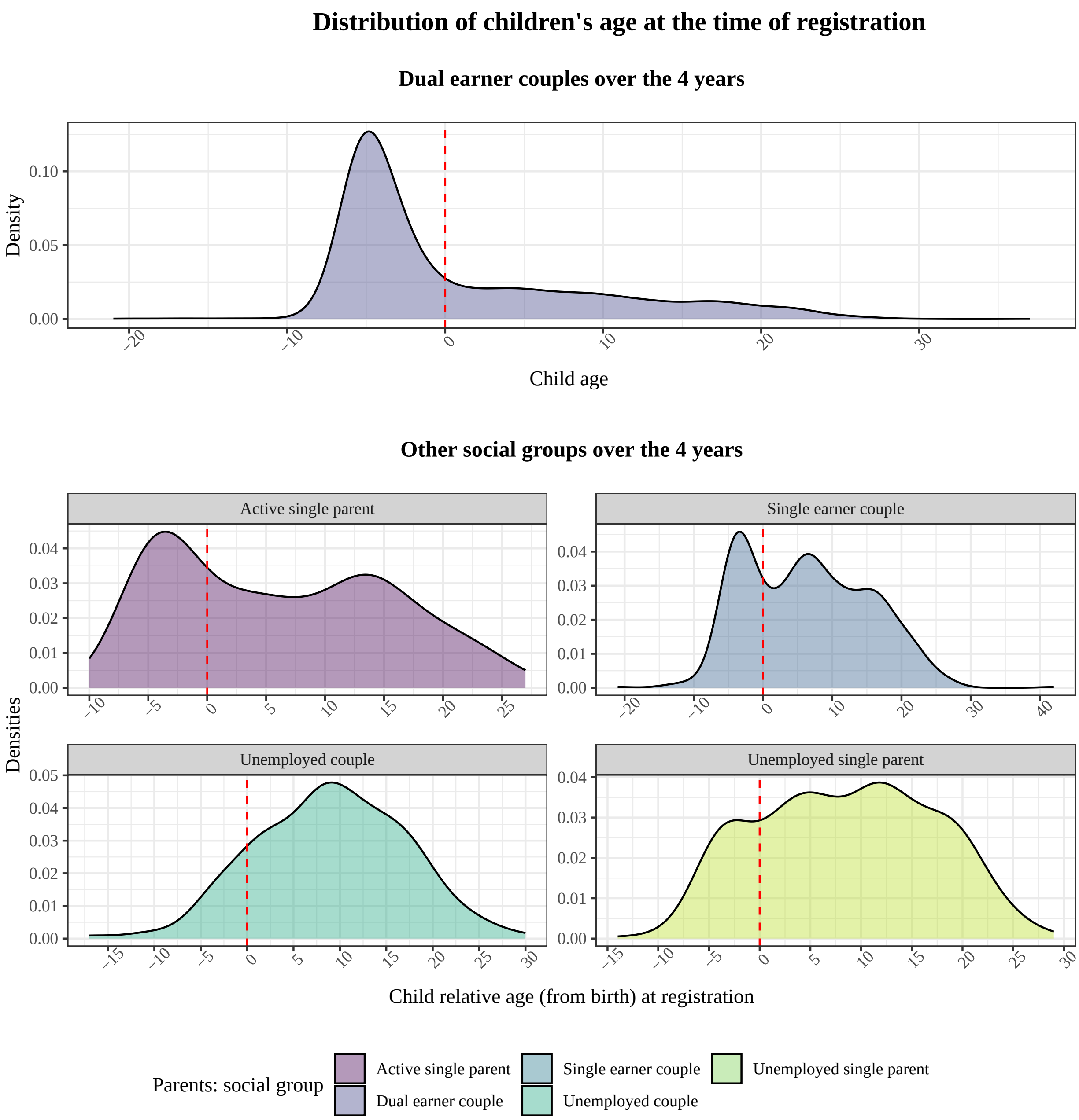
2. **Being able to wait**: Between birth and the assignment committee, parents require other childcare arrangements (including parental care) and wait for the committee's decision, with no guarantee of placement even when they registered early.

[62] Time since registration, residence, family's situation and employment and relationship status together

[63] Throughout every communication regarding daycare assignment procedures, parents are advised to contact the relevant administration *"as soon as their pregnancy is officially declared"*.

Indeed, for parents who anticipate their childcare needs, early registration clearly emerges as the dominant strategy, if they can afford waiting. Additionally, early registration during pregnancy reduces the needs (and costs) for childcare during the waiting period for the committee's response. Our sample is self-selected and if parents follow this dominant strategy, there should be a large bunching of registrations around 5 months before birth. This is precisely what we observe in the top panel of Figure 1.9 for dual-earners.

Figure 1.9: Strategic timing of registration of dual earners



Sources: ISAJE, Case Study I − 2020 : 2023 − first rounds only.
notes: Top panel present kernel density estimates of children age at time of registration and expected entry in september for dual earner couples.
Bottom panels present kernel density estimates of children age at time of registration for other social groups. For every estimates:
− One observation is kept for each child presented to one or other of the two committees,
− The variable analysed is constructed as the difference, rounded in months,
between the date of birth (actual or expected) and the date of registration.

A stable couple's privilege Parents in employment usually have to think and decide on their childcare arrangement early, sometimes even before conception. For dual earner couples choosing daycare, early registration is a dominant strategy and parents are incentivised to do so. These parents register young children for an early entry and this result is neither surprising nor negative *per se*.

However, this is a stable couple's privilege, as one can see in the distribution for other social groups in the bottom panels of Figure 1.9 and the empirical cumulative distributions in Figure D.24 in the Appendix. Over 60% of dual earners are registered before the birth, and over 50% registered more than three months before birth. This is more than twice as much as single-earner couples or single mothers in employment, for whom there is only a small bump of early registration.

This is pure endogenous selection at play as we only observe registered parents, and not parents who did not register. There are two obvious explanations for these differences: on the one hand, one parent may chose to care for their child themselves for some time and use childcare later. On the other hand, couples and jobs are unstable and break-ups and lay-offs are more likely to occur after the 3rd month of pregnancy. Both affect childcare affordability, needs, and ability to wait for the committee's decision. These parents cannot play the same strategy as dual earner couples and with born children, waiting is more expensive as it require alternative childcare arrangement. Moreover, they register older children and are competing over other buckets. Those have fewer available seats and so parents of older children face worse odds. The weights of social groups are highest for single parents working, but dual earner couples only have 1 point less. For them, registration a quarter earlier is enough to have higher priorities than single parents working. After 2022, single parents with no job need to be registered one year earlier than dual earner couples to reach their baseline priorities. Same for single earner couples and even longer for couples without jobs.

In the end, these weights create large inequalities of opportunities by allocating more seats for younger children mostly demanded by dual earner couples, giving this group the second highest social weight while it is the largest share of demands, and the opportunity to largely influence their priorities by applying as soon as they know of the pregnancy. This strategy is only accessible to stable families: stable couples, with stable jobs, who knew about their childcare needs and when to register, most likely before getting pregnant. For parents that *cannot* decide that early, they compete over fewer seats, opportunities are lower and furthermore, social weights strongly penalise parents whose childcare needs may be related to the events that lead them to be single-earner couple or single parent, and not dual earner couple. Social weights recognise the inability to wait of working single parents but the latters are still hurt by time since registration, as we have previously shown. The low weights for unemployed parents makes it very hard for them to access daycare. Furthermore, their job opportunities may involve atypical, flexible and unpredictable working hours that can be hard to reconcile with childcare services with typical opening hours. Allowing registration from the pregnancy declaration may have seemed like a harmless choice but it shapes the composition of the demand. It creates strong inequalities of opportunities for parents who *cannot* decide that early : in particular precarious couples with precarious jobs. The work of Pavolini and Van Lancker (2018) on the *Matthew effect* conclude from cross country analysis that structural constraints in childcare provision matter everywhere and tend to limit the uptake of childcare especially for children growing up in disadvantaged circumstances. We come back to this point in the discussion.

A) Other consequences of time since registration

With the previous results, we underlined that the majority of the largest group strategically register at the time of pregnancy to have higher priorities and that waiting is costly, especially for parents who decided on their childcare arrangement after the child is born. We show that the first also creates inequalities of opportunities even for strategic parents and generate variations in assignment probabilities by month of birth. The second requires parents to have outside options and paradoxically, increase inefficiencies by having more high priority parents choosing their outside options to the proposed assignment.

More inequalities of opportunities Even among dual earner-couples, these design choices generate negative consequences at equilibrium. The majority of demands are double earners and most play the dominant strategy. However, pregnancies occur throughout the year, and the *early-registration premium* decreases with pregnancies declared closer to the committee. As a result, parents who are well-synchronized have the opportunity to increase their priorities, which is not feasible for those whose pregnancy is declared closer to the allocation committee, unless they can wait until the following year. In the end, there are inequalities of opportunities even for strategic parents.

We discuss two situation in Example 1.4 in the Appendix. Well synchronised pregnancies offer parents with early childare needs an unfair advantage compared with parents with similar needs but whose pregnancy occurs at another moment. The same strategy does not yield the same "*return*" for parents whose children are born at different times of the year, or at a higher cost through a longer wait. And in this case, the child will be in a different age group, thus other buckets. In this market, daycare centres offer fewer seats for older children. Early registration priority gains can therefore be offset by capacity constraints at the bucket level, and for these very "*patient*" families, the outcome is still very uncertain. At equilibrium, these are inequalities of opportunities resulting from initial early-registration incentives which correlate assignment probabilities with pregnancy declaration and thus, month of birth. Figure D.28 in the Appendix displays the conditional assignment probabilities and 95% confidence intervals by registration and birth months[64]. The figure also provides a histogram and the admission status of children admitted on each date on the right scale.

There are substantial variations in admission probabilities by registration date, shifting from approximately 1/3 for registrations in the spring to 2/3 for registrations during the summer months (Example 1.4). Unsurprisingly, there is a negative correlation with the number of applicants. Admission probabilities also fluctuate with the child's birth date, with a pattern staggered by 5-6 months and of slightly lesser magnitude. The number of demands from children born in different months remains relatively constant from January to August. However, there are fewer demands from children born in the last quarter, with the associated assignment probabilities being the highest. For these parents, the declaration of pregnancy, occurring approximately 3 months before the summer, coincides with a notably reduced number of registrations. It is plausible that summer vacations postpones registration or makes them harder if offices close, compelling some to postpone their registration until September or even forgo it. Moreover, parents whose child is born in the last quarter of the year have to wait the most before getting a slot which may deter some parents from choosing daycare centres, as suggested by the lower number of registration.

Consequences: Unraveling The associations between registration months, birth dates, and admission probabilities are inherent consequences of the overarching market organisation. However, the early childhood department has informed us of increasing challenges, including a rise in refusals, numerous changes between requested days and actually signed contracts, and more. We have no data to measure this phenomenon yet, but the early childhood department currently considers this problem sufficient to motivate structural adjustment of the DAM. The time since application criterion creates opportunities to manipulate the assignment probability but requires a rather long waiting and without insurance of receiving a slot. Parents have to find alternative childcare solutions while they wait for a slot (including caring for their children themselves) and in case of rejection. Consequently, families enter the market with more or less outside options. This situation means that at the time of the committee, the list of preferences does not reflect the complete preferences of families. Some have outside options, the subjective rank of which is unknown in relation to the preferences expressed. If the committee offers a daycare that parents consider less preferable than their outside option, they will certainly refuse it. The problem is that priorities favour oldest

[64] They are calculated over children as principal statistical units and for all the rounds they could participate in, from 2020 to 2023.

demands, those whose preferences, situation or needs are most likely to have changed. Thus, because of outside options and preference updates, some with assignment probability are also the most likely to refuse them. In the current context, where families must confirm their registration shortly before the committee, many only withdraw after receiving an offer. This information asymmetry on outside options therefore causes part of the market to unravel[65]. In other words, this organisation results in:

1) Inflated demand;
2) Exacerbated congestion;
3) Seats being assigned to families whose childcare needs are already met – or can (more) easily be – rendering them more likely to withdraw their applications;
4) Over-rejection of single parents.

Although the organisation of a second round aims at partially rectifying this inefficiency, there is indeed a negative externality of withdrawing families on others. This situation underscores the importance of fostering sincere preferences and questions the relevance of *envy-freeness* as a stability criteria when envy is based on such inefficient and unfair rules.

V.4 Heterogeneous relative supply across age groups: lessons from Case study II

We saw in the subsection III.2 of section III that this marketplace features a sharp discontinuity in the total number of vacancies open for children born around January 1st. This DAM uses a simplified version of the problem and we implement the SPDA algorithm with sharp diversity constraints by birth dates. We want to estimate the effect of the sharp and common definition of the higher age bound of buckets and associated drop in number of open seats on assignment probabilities. We use a regression discontinuity design using time (in days) between January 1st 2023 and the child birthdate as forcing variable and analyse the jump in assignment probabilities and scores. We follow standard practice and let binwidth selection and polynomial order be data driven following Matias D. Cattaneo and Titiunik (2022). We also run a McCrary (2008) test of continuous densities of birth dates around the cut-off and present the results in Figure E.29 in the Appendix. All models are estimated using the *'rdrobust'* package.

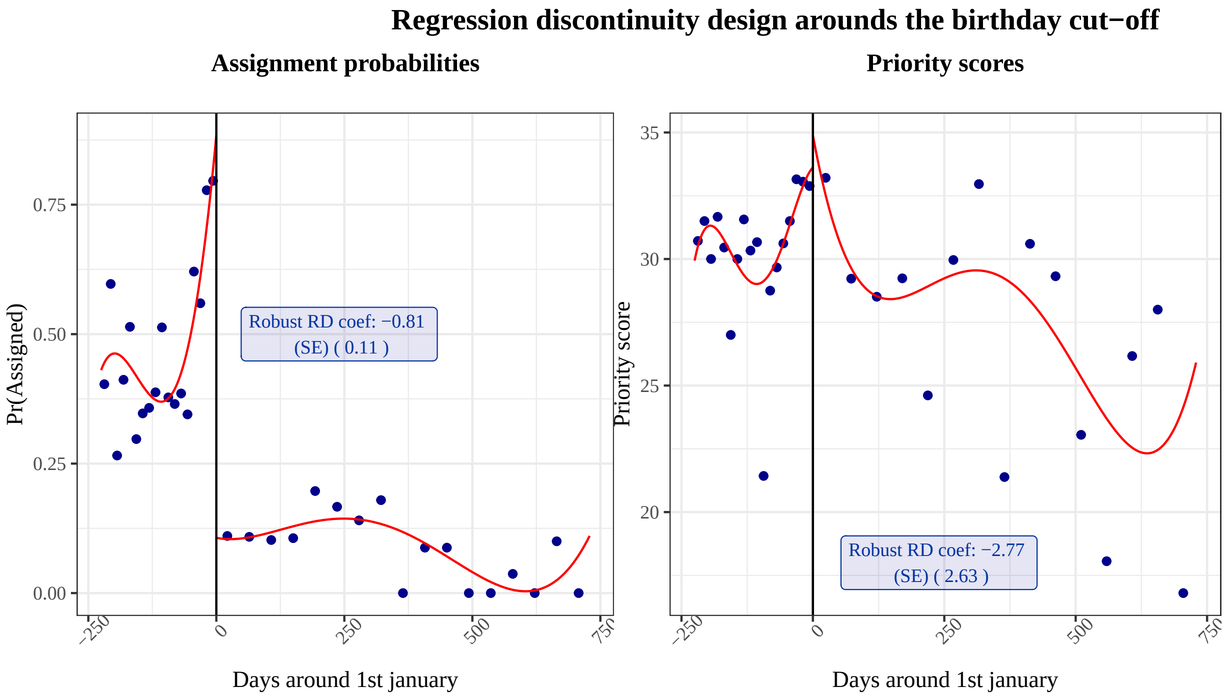
The main results are presented in Figure 1.10, with the discontinuity in assignment probability on the left and on priority scores on the right. As expected, the shortage of open seats for children born the calendar year before creates a sharp discontinuity in assignment probabilities. However, the priorities are not discontinuous around the threshold.

Interpretations The algorithm follows the definition of stability with diversity constraints and Theorem 2 imply that this is the best assignment for parents among those that respect priorities and diversity constraints. The matching is stable with diversity constraints and the *fairness* interpretation depends on their precise definition. Theoretically, those are the elements $(s_b, \succeq_b, c_b)$ that define the bucket structure and they are part of the definition of the problem. Although capacity constraints depend on actual physical constraints, buckets are strategic for both policymakers and daycare providers. The inequalities in assignment probabilities by date of birth only stems from a repartition of capacities across age groups that does not match the age structure of the demand. Policy makers knew that the relative size of age group mattered but did not realise parents differing only by a few days between their children's birth could face such unequal odds. The definitions of buckets are no less important than priority scores for shaping the distribution of assignment and generating inequalities. Priority weights change assignment probabilities across parents with different characteristics and are, by definition, intentional inequalities following *distributional justice* concerns. We showed they can cause *unfair inequalities* when groups with high social weights (single parents) are penalised by parents using a strategy (registration at birth declaration) they cannot use. Policymakers also control the relative supply they provide to different groups.

Finally, our two case studies show that assignment probabilities among registered parents strongly depends on well synchronised birth, but through a completely different mechanism. In Case Study I, unequal access by birthdate was

[65] Unravelling refers to the phenomenon where offers to families are rejected, accelerated, and require rapid adjustments, among other complexities, see e.g. A. E. Roth (2018)

Figure 1.10: Large discontinuity in assignment probabilities, no discontinuity in score



Sources: ISAJE, 'Family' files, Case Study II.
Notes: Data driven regression discontinuity with evenly spaced binscatters using the rdrobust package.

caused by strategic early registration from the largest group who could increase their priorities. It results in larger assignment probabilities for children born in the last quarter, registered around 3 months after conception. Children born in the two first quarters have smaller assignment probabilities. In case study II, time since registration does not affect priorities but policymakers supply less seats for children born in the previous calendar year. Thus, the effects of birth dates are reversed. Assignment probabilities are higher for children born in the two first quarters and much smaller for children from the end of the year. These results point the importance of all elements of the DAM, and the crucial role of buckets' elements and agents strategies in shaping the distribution of assignments.

VI Discussions

We started this project to provide high quality evidence on the effects of accessing daycare and for that, the first ordeal was to convince cities to let us randomise daycare assignments. To circumvent political resistance and legitimate ethical concerns, we learnt how committees assigned daycare slots. We asked how to define daycare marketplaces and the sort of optimality criteria we could define, and analysed the consequences of different design. We noticed similarities and contrasts with the school choice literature (Ehlers et al. 2014), and extend results from Kamada and Kojima (2023) and Delacrétaz, Kominers, and Teytelboym (2023) to define mechanisms tailored to each version of the problem we encountered. We then implemented these algorithms in 9 large urban district in France for up to four years and used data from two of them. We now discuss what we learnt, what these findings can tell about inequalities in access to formal childcare, and important policy implications.

How well do our model perform in the field ? One thing that all markets challenge participants to do is to decide what they like and one of our main contribution is to offer policymakers a *menu* of definitions and mechanisms to build daycare assignment marketplaces. Our work emphasises the trade-offs between different goals and lets them know the theoretical implications of their choices. Once policymakers define their DAM, the proposed procedures offer several distinct advantages over current practices in the field.

1) These computer-based tools can swiftly process a large number of applications, resulting in time savings for those involved in organising and participating in the assignment committees.
2) Our models make clear what can and cannot be done and the trade-off of each assignment mechanism. From a normative perspective, they ensure *fair* assignments in the sense that they eliminate *justified envy* based on a clear and chosen definition. Policymakers choose whether they want to strictly respect priorities, or accommodate more children by allowing weak accommodation and/or only consider initially feasible demands.
3) All assignment decisions are traceable and can be explained to every families. For each assignment we provide a table for every daycare that justifies every decision[66]. Transparency and accountability can be achieved, if policymakers choose to do so.
4) By ensuring that priorities are always respected and individually justifiable, they can create a public good[67]: they ensure *procedural justice* and can move ethical and justice concerns to *distributive justice*. Since priorities are always respected, society can debate on what constitutes fair (priorities) inequalities. For that, one requires *informational justice* i.e. that all participants have access, know and understand the rules.
5) Policymakers can define target assignment distributions by choosing the shares of seats they want to give to certain groups. Our models are flexible enough to consider many constraints. In particular, the definition of (soft) diversity quotas is a powerful tool to ensure that the assignment satisfies some distributional requirements. Without our procedure, that kind of policy objectives is very hard to achieve. Or at least, hard to justify.
6) These tools can also be used to simulate the effects of changing some parameters and inform policymakers using *what if* scenarii.

Good designs do not ensure well functionning markets In the first case study, we match demands over weekdays based on strong envy-freeness on initially feasible demands over flexible age groups and ranked through a priority score. The latter is made of three components: i) A disadvantage for demands outside the area, ii) social weights over groups based on the joint distribution of employment and relationship status summed with few additional weighted socio-demographic situations, and iii) time since registration, open from pregnancy declaration. The latter encourages early registrations and explains most variations in priorities. The majority group - dual earner couples - often require childcare arrangement earlier and most strategically register as early as possible. Using simulation of counterfactual assignment with alternative priorities, we showed that weighting social groups can effectively change assignment probabilities. However, allowing parents to manipulate their priorities through early registration strongly penalises single parents.

This is in part because early registration cannot be a strategy for those who became single parents later in their pregnancy. These inequalities are further aggravated by high social weights for dual earner couples who can anticipate more, wait longer and at a lower cost. Over the four years, single parents and couples where only one parent works have only seen their situation worsen. Strategic timing of registration does not have the same return when pregnancies are declared later in the year and they create inequalities of opportunities even among strategic parents through large variations in assignment probabilities by date of birth. In the second case study, the distribution of capacities across age groups is highly unbalanced and far from the corresponding demand. The sharp discontinuity in assignment probabilities by date of birth is another illustration of inequalities of opportunities stemming from the bucket structure.

[66] These tables are built from the database used to make the figure 1.5, add the initial capacities (by day) and cumulative use.

[67] According to Colquitt et al. (2001), justice perceptions can be broken down into several heuristics. Procedural justice concerns the processes, logic and deliberation behind a decision. Distributive justice concerns the allocation of positive and negative outcomes in a decision context and whether they are distributed equitably or deservedly amongst the affected population given their circumstances, performance or contributions. Interactional justice concerns the extent to which the affected individual is treated with dignity and respect by the decisionmakers. Finally, informational justice pertains to the information and explanations provided for decisions: are they candid, thorough, and tailored to individual needs?

A daycare assignment marketplace, like any market, is a social construct influenced by various factors that can either facilitate smooth transactions or give rise to frictions, inefficiencies, selection biases, and more. The choices made by policymakers matter in every aspect. Priority scores is not their only tool but it is a powerful one to enforce higher assignment probabilities for those with valued characteristics. However, buckets are even more crucial because they define who can access which capacities, how much capacities they receive and how to chose applicants. These buckets are part of the definition of the stability notion and as such, they should be defined with care.

Can daycare assignment be fair ? What would it means ? The trade-off between fairness and wastefulness is an inherent aspect of the daycare market. Our work offers policymakers different options regarding the concept of "*fairness*", enabling them to mitigate wastefulness without compromising a robust definition of envy-freeness. However, this definition depends on many parameters that affect the entire market structure. Inequalities of opportunity emerge at equilibrium and may be hard to justify when the definition of priorities is not tied to some notion of social justice. For instance, rejecting a set of families because one family registered early and cannot be served is arguably more contentious than when families are declined because it was not possible to accommodate a disadvantaged child.

Sometimes the problem is more about the *perception of fairness* rather than *fairness* as defined by absence of justifiable envy (Bullock, Williams, and Limbert 2003; Van Lancker and Ghysels 2016; Hufe, Kanbur, and Peichl 2020; Trump 2020). Perceived levels of justice are separate from purely self-serving rationalisations of a decision outcome. The very existence of priorities means accepting some have more chances of acceptance than others. It is a sensitive political question. However, most municipalities do not elicit preferences from their population to build priorities and face little accountability. The agreement between policymakers and daycare providers on the definition of buckets is central and yet, its strategic and fairness implications are absent from the debate. Even when criteria and weights are well defined and reflect a political objective, few are accessible and known to parents and in most cases, criteria are presented without their weights.

The instructions that accompany a mechanism (e.g., priority rules or strategy-proofness) are an important part of its design (A. E. Roth 2018). For instance, Hassidim, Romm, and Shorrer (2018) identify a substantial fraction of applicants in the Israeli Psychology Master's Match who misrepresent their preferences even though the mechanism is SPDA, which is strategy-proof. They further show that academically weaker applicants are more likely to misunderstand the instructions provided by the clearinghouse, explaining why they are more likely to misreport their preferences. Traditional incentives of more well-off families for using formal care (i.e., dual careers and few informal care opportunities) are often less salient for underprivileged families. It may thus be necessary to emphasise other motives and make registration safe. If parents intend to enroll their children, the lack of transparency may hinder the parents' action and make the enrollment process too costly (Carbuccia, Thouzeau, et al. 2023).

The choice of childcare among other strategies Securing a slot strongly depends on institutional settings such as priorities in daycare assignment mechanisms, congestion, access to alternative options and more. Importantly, some parents react to the market structure, but not all of them, and sometimes at the expense of more vulnerable families. There are several similar results in the school choice literature, starting with strategic preferences reporting and the externalities on sincere families, usually from lower economic background (Pathak and Sönmez 2008). Strategies of more affluent families are not limited to these. Fack and Grenet (2010a) showed that high quality schools raise housing prices in Paris in the previous residence-based assignment system. A very recent working paper by Bjerre-Nielsen et al. (2023) showed that the introduction of a residence-based admission criteria in the school choice system in Denmark doubled address changes and only in areas where the incentive to manipulate is high-powered. In the childcare market, Steinberg and Kleinert (2022) analyse the timing of registration for childcare testing rational choice theory and find consistent estimates with their theoretical predictions. In this setting, *bunchers* are optimising and their strategies are not limited to the timing of registration. Other decisions or outcomes are related. For instance, Bernal and Keane (2010) estimate quasi-structural models of the joint decision of childcare choice and labour market participation to measure the effects on children development. Their estimates focus on single mothers in the USA and use the Welfare reform of 1996 to instrument childcare use. Their results show large negative effects of the joint decision to work and use childcare on children cognition. Allègre, Simonnet, and Sofer (2015) model simultaneously labour market participation and type of childcare choice using a selection model *à la* J. Heckman (1990) and estimate a multivariate probit on French data from 2002. Their results show that monetary incentives matter for chidcare use, but not which childcare is used. They also show that tax-rate matters only on labour market participation at the intensive margin. However this model lacks clear identification and omits important variables. In France, daycare centres have an objective comparative advantage in that they are, in general, the most affordable childcare solution. They also have a shared perceived comparative advantage over children socialisation, and researchers have documented that parents prefer collective childcare to childminders (Bouteillec, Kandil, and Solaz 2014b; Cartier et al. 2017 ; Caenen and Virot 2023). Pora (2020) uses large but staggered increases in daycare provision across municipalities and time to estimate the effect of these apparent supply shocks on mothers' labour market participation. On average, he finds no effect on labour market participation but documents a crowding-out effect on registered childminders. These results suggest a general equilibrium effect in which women attached to the labour market shifts from childminders to daycare and manage to secure access for their children. In our setting, parents who work are much more likely to be seated because of priority rules and strategic behaviour, which supports the general equilibrium hypothesis and results of Pora (2020).

Surprisingly, this simple supply/demand/equilibrium reasoning is very much overlooked. However, it helps to think of how policymakers may improve welfare and where interventions are likely to make things worse. There may also be *optimisation frictions*: tedious application processes and complex tax subsidies schemes, among others, which may lead parents to perceive childcare as out of reach (Ünver, Bircan, and Nicaise 2018; Weixler et al. 2020a; Hermes et al. 2021, 2022; Valant and Weixler 2022). Policymakers, daycare providers and parents have to make sequences of nested decisions regarding childcare and every choice has an opportunity cost. In particular, families from disadvantaged backgrounds face greater challenges in securing a place for their children in high-quality daycare centres, limiting their access to crucial developmental opportunities (Gambaro, Stewart, and Waldfogel 2015). A crucial question is: how large is the opportunity cost of the Matthew effect ? $I.e.$ what is the welfare loss due to "*the observation that the benefits of government spending on social policy disproportionally accrue to middle- and upper-class relative to other social groups*" ? (Pavolini and Van Lancker 2018)

The Matthew effect as a political choice Single parent families hold a pivotal place in the discourse surrounding social investment. Policies for reconciling work and family life, such as formal childcare arrangements, aim to both enable parents to work and contribute to children's healthy development, ultimately reducing social inequalities from childhood (Van Lancker 2013). We showed that in these two settings, policymakers strongly influence the assignment distribution and favour more affluent households and early entry in childcare. Time since registration negatively affects single parents and unemployed parents. The current priority system strongly advantages stable families: stable couples, with stable jobs, who knew about their childcare needs and when to register, most likely before getting pregnant.

If we think of the ambitious goals for early childcare policies in the social investment narratives (Morel, Palier, and Palme 2012b; A. Hemerijck and Huguenot-Noël 2022), this sort of assignment distribution is very regressive. Two recent meta-analysis show that access to childcare services have large benefits for children and the positive effects are primarily concentrated among children from socio-economically disadvantaged backgrounds (van Huizen and Plantenga 2018; Schmutz 2024).

These results have strong economic and fairness implications. From an economic perspective, the Matthew effect creates large efficiency losses by reduces crucial development opportunities for low-income families. Another way of thinking of this problem is to think of the matching surplus and opportunity costs such priority generates for parents, policymakers, and society. A first important thing to note is that current weights generate actual private benefits while opportunity costs are mostly uncertain and in the long run. Neimanns (2022) argue that there are short-term electoral gains to favour more affluent families and in practice, they do favour more affluent families. Moreover, fees increase with incomes so it is also cheaper to favour them. Affluent families also benefit from collective childcare who are in general cheaper than childminders or private childcare services. Their career may be less affected by the *child penalty* $i.e.$ the gender gap in the labour market after the birth of the first child[68].

When policymakers favour working parents, it acts both as a reward for hard working and deserving mothers and a mean of reducing the child penalty. However, the counterfactual for these parents may not have been parental care

[68] The seminal work of H. Kleven, Landais, and Søgaard (2019) analyses career trajectories using retirement data in Denmark and found a persistent effect over 20 years for women, with an approximately 20% decrease in employment, income, and other indicators. There is no penalty for men, but for women, the loss is more pronounced in families with traditional values.

because these parents have other outside options. Conversely, low-income families who do not get a seat usually cannot afford alternative formal childcare arrangement. This in turn may be responsible for larger child penalty at the bottom of the income distribution. Meurs and Pora (2020) use an event-study design *à la* H. Kleven, Landais, and Søgaard (2019) to estimate the *child penalty* In France. They find a persistent 20% decrease in labour market participation over 10 years, a 40% decrease in income, a 10% decrease in hours worked, and a declining trend in wage rates for women following childbirth. For men, there is no economic penalty associated with the arrival of a child; in fact, their incomes increase in the years following childbirth[69]. An innovative aspect of this research is the analysis of gender penalty by income deciles. Their results indicate that the penalty is much more pronounced and persistent at the lower end of the income distribution, whereas it disappears completely for women at the higher end of the distribution. Bazen, Xavier, and Périvier (2022) estimate similar models focusing on cohorts of new entrants in the labour market and find that the child penalty is much higher for mothers with lower education. The latter group is disproportionately represented among single mothers. Last, because of gender norms and marital specialisation, Bonnet, Garbinti, and Solaz (2021) document large gender gaps in standard of living after divorce and massive labour market re-entry by previously inactive women. However, access to childcare is often a necessary condition to look for, and eventually get and hold a job (Gorey 2009). Labour market effects also depends on job opportunity and market tightness, travel distance and institutions such as flexible hours (Flèche, Lepinteur, and Powdthavee 2020; Le Barbanchon, Rathelot, and Roulet 2020; Chung 2020). Together, these evidence show that the Matthew effects largely depends on political choices and its cost on society is largely carried by women an children in low income groups.

Can we scale-up ? Should we ? An ambiguity between *market design* and *computer science* lays in the way users or policymakers think of the role of such mechanism. In our fieldwork, some municipalities were interested in our algorithm because they reduce the amount of back-office work and considered it as an updated IT system. Others valued the envy-free and transparency properties of our designs and even communicated lottery realisations to parents to justify assignments. In general though, the automation occurred with rather minimal information provided to families, primarily to preserve the integrity of the research protocol[70] and avoid political backlash. As we have seen, mechanisms are appropriate for DAM based on our models of economic agents. We understand they can react to different designs or features. The official adoption of our tools will likely elicit different reactions from families, early childhood professionals and policymakers.

Local power dynamics are important and incumbents value their autonomy. Even the inter ministerial mission assigned to an elected member of the Association of French Mayor (AMF) could not provide recommendation on *which* criteria *could/should* be used for priorities (Laithier 2018). Their political opposition has been made quite clear. After publishing this report encouraging transparency, meeting us several times, AMF sent a letter[71] to every mayors part of the early-childhood group deterring them from joining our project. Our approach was both seen as a threat to the self-government of territories and was amalgamated with IT tools:[72] "*This approach, which consists of replacing the human dimension and the finesse of the work carried out by elected representatives, in particular in the committee responsible for allocating daycare seats, with an algorithm, is in total contradiction with Laithier (2018).*" and "*The AMF's fear is that the research carried out by the CNAF will eventually lead to the generalisation of the algorithm for allocating crèche places imposed by the family branch on managers of early childhood education establishments (EAJE) or, at the very least, that co-financing will be conditional on its use*."

The political aspect of this research is not secondary. We changed market structures. We are accountable for which family received a slot and which did not. To respect ethics and scientific integrity, our work was discussed and approved by an international scientific committee with researchers from different fields (sociology, economics, education science and so on). As researchers, we must acknowledge our responsibility and part of it is to report what we learnt. While we faced resistance in the early stage, we received a lot of positive feedbacks from those we worked

[69] also see the recent publication by France Stratégie signed by Dherbecourt and Flamand (2023)
[70] Parents were informed that the city hall took part in a research on early childcare as bare minimum.
[71] See the link on their website : https://www.amf.asso.fr/
[72] Original text is : "Cette démarche consistant à remplacer la dimension humaine et la finesse du travail réalisé par les élus, notamment en commission d'attribution des places en crèche, par un algorithme est en totale contradiction avec le vade-mecum" and "La crainte de l'AMF est qu'à terme la recherche menée par la CNAF aboutisse à la généralisation de l'algorithme pour l'attribution des places en crèche imposée par la branche famille aux gestionnaires d'établissements d'accueil du jeune enfant (EAJE) ou, à tout le moins, que des co-financements soient conditionnés à son utilisation."

with, such as informal confirmations of reduced interventions from elected officials or less formal complaints from families.

Depending on the quality of the tool offered, the adoption of technology can also lead to endogenous changes in committee organisation systems. Many territories opt for holding 1 to 4 allocation committees, often constrained by logistical issues that would be immediately addressed with a suitable tool. The problem would then be quite different from how we have conceptualised it. Dynamic assignments are fairly distinct theoretical problems[73]. An important feature that guided our models and assignment mechanisms is that capacities are mostly released at once, in September, and then, organising the convergence of supply and demand at a specific time allows for more choices for parents and managers. Market thickness offers participants a public good through larger *menus* for both sides of the market but as A. E. Roth (2015) points out, "*Part of making a market thick involves finding a time at which lots of people will participate at the same time. But gaming the system when the system is "first come, first served" can mean contriving to be earlier than your competitors.*" These strategic players also are those with outside options which may cause the market to unravel. Family choices are heavily driven by the existence of multiple outside options for care and the results on school choice and outside options are scarce (Akbarpour et al. 2022). Recent evidence show that families do have heterogeneous beliefs about childcare access and quality (Boneva et al. 2022). One central question is to know whether participation is driven by preferences *per se* or misbeliefs of certain types of agents (e.g., low income families). Currently, there are no studies on the impact of misbeliefs of parents for participation decision into a school choice type of market. Finally, these models are partial equilibria and the actual participants' strategy may involve other choice sets outside of this market, dynamic interactions with other markets or itself (when families remain for more than a year). Parents may have preferences over a larger set of options that are not represented in this market (or any market). While this has little impact on our modelling assumptions to design assignment mechanisms, it has tremendous implications in real life settings, especially when outside options are not accessible to all parents and a market may be noxious because some of the transacting parties are vulnerable (Li 2017a).

Scaling up creates new challenges. Details matter and designing well functioning marketplaces may require inputs from different types of agents. We think scaling-up should involve public institutions (Municipalities, CAF, PMI,…) and citizens, researchers and engineers, so that the proposed services perform *well* based on normative criteria policymakers should define. Early-childhood education and care is seen as one of the most important policies to reduce inequalities and more. This provides clear objectives and we can define measures to assess whether daycare assignment mechanisms do well. Our work would benefit from the expertise of researchers from other related disciplines, such as organisational sociologists to investigate needs, reluctances, and reactions among parents, policymakers and childcare providers. Eliciting preferences to define *fair* priority rules would require the involvement of psychologists, behavioural scientists and sociologists. Fostering access to daycare for certain population may require additional interventions which could involve associations and other local organisations.

In the end, these daycare assignment mechanisms can open a path towards easier and more transparent assignment mechanisms, and provide new tools to define distributional objectives. It definitely opens new research agendas by generating new data that so far, have not often been accessible to researchers. The embedded lotteries can help answer important policy relevant research questions. For instance, unravelling causes many inefficiencies and adjustment but we can identify compliers characteristics, at least at the margin. Ultimately, we have enough random variation across all marketplaces we worked with to create a sample of about 20 000 demands among which 3 000 were subject to random assignments. Our next work will be to assess the effect of accessing daycare on parents' labour market decisions and marital stability.

[73] See *e.g.* Kennes, Monte, and Tumennasan (2014) and Grenet, He, and Kübler (2022)

Appendix

A A short description of the French family benefits system for parents of young children

A.I Types of childcare available in France

Daycare centres are usually public or strongly depend on public funds but there are some private for-profit centres and other arrangements. Here are the different types commonly found:

- **Public Childcare Centres** (*Crèches*): Public childcare centres, often referred to as ”*crèches*,” are operated and funded by local municipalities or public institutions. They provide full-day care for children, typically between the ages of three months and three years. Crèches are regulated and staffed by qualified professionals. There are laws on the number of children by adults that vary according to children's ability to walk and binding regulations on care practices, nutrition, environment,... They are mostly publicly funded and parents' fee depend on their income and family situation.
- **Private Childcare Centres**: Private childcare centres are operated by independent organisations or individuals. These centres often offer similar services to public crèches, but they may have different pricing structures or cater to specific age groups. Moreover, some employers can either build their own daycare centres for their staff (and sometimes accept applicants from non-staff parents) or ”buy” seats in other centres. These often exists in some public administrations (hospital and universities) and large private firms.
- **Micro-Crèches**: Micro-crèches are small-scale childcare centres that typically accommodate a limited number of children, usually up to 10. These centres provide a more intimate and home-like environment with a focus on individualised care.
- **Childminders** (*Assistantes Maternelles*): Chidlminders, known as ”*assistantes maternelles*,” are registered and licensed childminders who offer childcare services in their own homes. They care for up to four children, providing nurturing and personalised settings. They are directly hired by parents who receive subsidies (CMG, see below) and tax credits. Some are also hired by city-halls
- **Parent-Run Co-Operatives**: Parent-run co-operatives are childcare arrangements where a group of parents collaboratively organises and manages a childcare facility. Parents are involved in the decision making and operation of the cooperative, ensuring a high level of parental involvement.
- **Nannies** (*Nounous*): Nannies are private childcare providers who directly work for individual families. They provide personalised care for children in their families' homes and offer flexible scheduling to meet the family's specific needs. Nannies are typically employed by the family and may live-in or work on a part-time basis.

A.II Aids directed specifically to parents with children under 3

In France, PAJE (*Préstation d'Accueil du Jeune Enfant*) is a comprehensive family benefit programme aimed at supporting parents in the early stages of their child's life. PAJE provides financial assistance and benefits to families with children under the age of three. The programme includes several components:

- *Prime à la Naissance*: the premium at birth or on adoption is paid, depending on income[74], on the birth of a child, or at the time of the adoption of a child under 20 years old. For eligible households, it amounts to € 1 019.40 paid on the seventh pregnancy month or at the time of adoption.

[74] In 2023, for single-earner couples, the eligibility threshold of the premium at birth is € 33 040 of yearly income 2 years earlier for the 1st child, € 39 648 for the second and € 47 578 for the third. For single parents or double-earner couples, these thresholds are respectively €43 665, €50 273 and € 58 203. For families with more than 3 children, thresholds increase by € 7 930 for every additional child, whatever the household situation.

Table A.2: Change in the main type of childcare for children under 3 on weekdays between 2002 and 2021

	2002	**2007**	**2013**	**2021**	**Counterfactual†: if parents had their first choice**
Parents	70	63	61	56	36
Grandparents or other family members	4	4	3	3	2
Childminder or Shared Childcare Facility (MAM)	13	18	19	20	23
Early Childhood Education and Care (EAJE)	9	10	13	18	35
Other forms of childcare arrangements[1]	4	5	5	3	4

[1] e.g. in-home care, school, shared childcare facility (MAM), friend, neighbor, babysitter, or other non-family member, kindergarten, after-school care, leisure centre, or specialised establishment.
† In the survey, the ask families what their first choice whould have been, regardless of what they are using now.
* *Note:* The week is counted from Monday to Friday, from 8 a.m. to 7 p.m.
** *Reading:* In 2021, 56% of children under 3 years old are primarily cared for by their parents from Monday to Friday, from 8 a.m. to 7 p.m., compared to 70% of children of the same age in 2002.
*** *Source*: DREES (French Directorate for Research, Studies, Evaluation, and Statistics), surveys on childcare arrangements for young children, Metropolitan France, children under 3 years old.

- *Allocation de Base (AB)*: the basic allowance follows the payment of the premium at birth or on adoption and takes two values (full or partial), depending on income[75]. It is paid from the month following the birth until the last day of the calendar month before the child's 3rd birthday (or 3 years from the month following the adoption, up to a maximum child age of 20).

- *Prestation partagée d'éducation de l'enfant, PreParE* The shared allowance for the children's education allows one or both parents to stop or reduce their activity to take care of their children under 3 years old (under 20 years old in case of adoption). For the first child, each parent can take up to 6 months up to the child first birthday or the full year for single parents. For families with more than two children, they can take up to 24 months each up to the youngest child's third birthday. When parents entirely stop working, they are eligible for a € 428,71 monthly allowance. If they work up to 50

- *Complément de libre choix du mode de garde (CMG)*: The free choice of activity supplement is paid to the household or person who directly employs someone to take care of a child who is under 6 years old or place the child in a micro-crèche.

The eligibility and specific details of each component of PAJE vary based on factors such as family income, the number of children, and the type of childcare chosen.

In addition to the PAJE program, there are several other aids available to parents with children under three in France. These include:

- Aide Personnalisée au Logement (Personalised Housing Assistance): This financial aid helps families with housing costs, including rent or mortgage payments. Eligibility and the amount of assistance depend on factors such as income, family composition, and housing expenses.

- Allocation de Rentrée Scolaire (Back-to-School Allowance): Although primarily aimed at school-aged children, this annual allowance provides financial assistance to low-income families to help cover the costs of school supplies and equipment. It can be applicable to families with children entering early childhood education.

[75] The full monthly payment of AB amount to € 184.81 and the partial payment is € 92.40 in 2023. Single earner couples with one child get the full payment if their yearly income is less than € 27 654, €33 185 for two children and €39 822 for three. The thresholds for partial AB are the same as those for the premium at birth. For single parents or dual-earner couples, the thresholds for the full AB are €36 546 for one, € 42077 for two and € 48 714 for three children. With more children, thresholds increase by €6 637 for each additional child.

- Allocations Familiales (Family Allowance): This universal benefit is provided to families with at least two children and helps support the general costs of raising children. The amount of the allowance varies based on the number of children and their ages.
- Aide au Temps Libre (Leisure Time Assistance): This benefit aims to facilitate the participation of children in recreational and cultural activities. It can help cover the costs of sports clubs, music lessons, cultural outings, and other extracurricular activities.
- Couverture Maladie Universelle Complémentaire (CMU-C) or Complémentaire Santé Solidaire (CSS): These are healthcare coverage programs that provide access to medical services, including for children under three. They offer assistance to low-income households who may not have adequate health insurance coverage.

B The ISAJE research project

General idea and constraints The research design of ISAJE relies on the implementation of an algorithm to allocate daycare seats to families and retrieve those who were subject to randomisation in the process. The share of demands subject to a lottery can be small relative to the whole population and strongly depends on the coarseness of the priority rules. We defined a set of constraints to search for cities. They had to

1. Use well defined priority rules that do not depend (or slightly depend) on tight criteria such as "date-since-application", or who were willing to give them up for the experiment;
2. Allocate daycare seats through a main allocation committee ;
3. Are willing to substitute an algorithm to their committee, and commit to its results in large;
4. Have a large number of demands so that the experimental sample is not too small;
5. Are significantly rationed, measured by the ratio of theoretical capacity over the total number of children under 3.

As a rule of thumb, we restricted potential prospects to local administrations whose theoretical capacity is above 600 seats. We started prospecting municipalities in 2019 and stopped in 2023. We worked with 9 large urban cities, some for several years. We are running another round of assignments in 2024 for a subset of cities who asked for this service to be continued.

Data limitations because of ongoing research Starting this project, our primary research objective was to analyse the impact of daycare access on families using counterfactuals derived from comparison groups within the same local labour market. To achieve this goal, we convinced 9 large urban centres to participate in this research project. The most important aspect for the evaluation was to ensure that enough assignments and rejections would be random and traceable, while limiting disruption in the ecological conditions. The models we present in section IV provide solutions compatible with existing practices, a necessary condition to gain support from incumbents without infringing on local policies. In practice, we operated from 2020 to 2024 as subcontractors so that city-halls remained data controllers, as per GDPR. We charged no fees and provided detailed feedbacks and policy recommendations based on data analysis. For some, we also ran simulations, showcasing another advantage of our tools: simulating assignments with counterfactual priority rules to inform policymakers of possible distributional consequences.

Because of our subcontractor status we cannot use all these data for research yet. However, Valence-Romans Agglo enabled us to keep fully anonymous datasets. Another agreed we could report results using their data without disclosing the source. Our analysis focuses on priorities and how parents react, and the role of age groups and diversity requirement.

B.I Marketing the ISAJE experiment

Figure B.11: Flyer to recruit new territories, P1

Isaje une recherche sur l'accueil du jeune enfant

Une ambition : adosser l'investissement social dans l'accueil du jeune enfant à une démarche scientifique

L'accueil des jeunes enfants est une politique déterminante pour la société française et mobilise de nombreux acteurs, au premier rang desquels se trouvent les collectivités territoriales et les Caisses d'Allocations familiales (Caf). Il participe à la réduction des inégalités femmes-hommes et, parce qu'il intervient à un moment clé du développement des enfants, est susceptible de réduire les inégalités d'acquisition de compétences liées à leur origine sociale.
Les acteurs publics jouent donc le rôle d'investisseurs sociaux. Ils s'intéressent au développement de l'enfant et anticipent que les efforts d'aujourd'hui généreront des gains économiques et éviteront des coûts sociaux futurs. La crédibilité de cette démarche repose sur l'existence de travaux scientifiques permettant de mesurer les effets des actions d'investissement social et les gains induits pour la société. C'est pourquoi nous lançons **le projet de recherche ISAJE pour « investissement social dans l'accueil du jeune enfant ».**

Le projet de recherche Isaje

Ses objectifs

L'objectif de cette recherche est de mesurer **les effets de l'obtention d'une place en établissement d'accueil du jeune enfant** (EAJE) sur trois dimensions :

- Le développement des enfants dans les compétences langagières, cognitives, motrices, socio-émotionnelles ;
- Les conditions de ressources des familles et l'activité professionnelle des parents ;
- Le bien-être subjectif des parents et le quotidien du ménage.

Elle s'attache en outre à **mesurer si ces impacts sont différents suivant l'origine sociale des enfants.**
L'originalité de cette recherche et son intérêt majeur résident dans sa capacité à identifier des relations « de cause à effet » entre le fait d'avoir accès à un mode d'accueil collectif et la situation de l'enfant et de sa famille. Les études permettant de montrer les effets des modes d'accueil sur le développement des enfants sont rares en France, et ce sont alors souvent quelques études emblématiques étrangères qui sont citées pour en justifier l'intérêt.

Un partenariat institutionnel et scientifique

Ce projet est porté par la **Caisse nationale des Allocations familiales** (Cnaf) avec le soutien de la **Délégation interministérielle à la lutte contre la pauvreté des enfants et des jeunes.**
La méthodologie de la recherche est conçue par deux chercheurs[1] de l'**École d'économie de Paris (PSE)** et de l'**École polytechnique** avec l'aide d'un conseil scientifique pluridisciplinaire et international composé de chercheurs, d'institutions et de l'équipe de recherche[2]. Sa gouvernance est assurée par un comité de pilotage dans lequel les territoires engagés dans la démarche sont représentés. Ce projet est proposé en étroite collaboration avec les Caf concernées.

[1] Arthur Heim (Économiste, doctorant à PSE) et Julien Combe (Économiste, École polytechnique).
[2] Les chercheurs sont : Sylvana Coté (Psychologue, University of Montreal), Orla Doyle (Économiste, University College Dublin), Denis Fougère (Économiste Sciences Po Paris), Claude Martin (Sociologue, EHESP Rennes), Yusuke Narita (Économiste, Yale University), Laurent Toulemon (Démographe, Ined Paris), les institutions représentées sont la Cnaf (Florence Thibault), la DREES (Emilie Raynaud), l'INJEP (Thibault de Saint-Pol / Aude Kerivel), et le projet est porté par Marc Gurgand (PSE), Arthur Heim (PSE-Cnaf), Julien Combe (École polytechnique), Jeanne Moeneclaey (Cnaf).

Figure B.12: Flyer to recruit new territories, P2

Comment ça marche ?

Le processus d'attribution des places

L'accueil d'un enfant en crèche est le résultat d'un processus complexe impliquant d'une part, des préférences et des contraintes du côté des familles, et du côté de l'offre, un nombre de places souvent limité, dont l'accès est défini par les critères de priorité établis localement. Ainsi, le processus d'attribution induit des différences de profils entre les enfants accueillis en crèche et ceux qui ne le sont pas. Dès lors la comparaison de ces deux groupes ne permet pas d'identifier les effets liés à l'accueil en lui-même. Le projet ISAJE vise à répondre à cet enjeu en organisant un suivi de familles statistiquement identiques dont une partie est accueillie en crèche et l'autre est en liste d'attente. Le projet porte sur des territoires marqués par un excès de demande de places en crèche par rapport à l'offre disponible, de sorte que la recherche n'implique en aucun cas de priver des enfants d'une place d'accueil.

Evidemment, chaque territoire a ses propres critères pour l'attribution des places et il faut tenir compte des préférences exprimées par les familles. Le projet de recherche ISAJE propose donc **une solution technique prenant en compte les souhaits des familles et les critères de priorisation des communes tout en permettant la réalisation de la recherche, qui suppose d'introduire de l'aléa dans l'attribution des places.** Cette solution concerne uniquement la commission d'attribution liée à la recherche (*a priori* printemps 2021).

Le projet de recherche mobilise une procédure automatique d'affectation utilisant des critères explicites et transparents, fournis par les territoires, qui garantit une stricte égalité de traitement à profil donné. Il accorde ainsi **aux plus prioritaires le meilleur choix possible parmi les vœux formulés faisables.** De plus, cette procédure permet de prendre en compte la complexité de la réalité : elle tient compte des demandes par jours spécifiques, peut introduire des places réservées (pour renforcer la mixité sociale par exemple), ou avoir un traitement adapté aux dossiers sensibles (situation de handicap par exemple).

Des enquêtes auprès des parents et des enfants

Une fois les attributions de places réalisées, **un échantillon de familles** bénéficiaires et non-bénéficiaires d'une place **sera suivi au cours des trois années suivantes.** Des enquêtes seront régulièrement menées auprès des parents et une évaluation des compétences des enfants à 3 ans sera réalisée par une ou un psychologue, psychomotricienne ou psychomotricien.

Valence-Romans agglo : un territoire pilote

En 2020, l'agglomération de Valence-Romans s'est portée volontaire pour participer au pilote de ce projet. La commission d'attribution du printemps 2020 portait sur 1014 demandes pour environ 380 places à temps plein réparties dans 43 structures (dont 12 assistantes maternelles), divisées en 5 groupes d'âges. Les familles exprimaient jusqu'à 8 vœux (3 en moyenne) avec des jours spécifiques. Les priorités des familles sont définies par un score construit à partir des caractéristiques sociodémographiques des familles auquel un numéro au hasard est ajouté de sorte à départager les dossiers ayant le même niveau de priorité.
En utilisant la procédure ISAJE, 400 familles ont vu leurs demandes satisfaites, soit 39 % des demandes. Parmi elles, 64 % ont obtenu leur premier choix, 17 % leur second. En comparaison avec ce qu'aurait fait la commission : 83 % de dossiers auraient eu une affectation identique. Une première enquête téléphonique sera réalisée au printemps 2021 auprès de l'ensemble des parents ayant demandé une place en crèche.

ISAJE a besoin de vous !

La recherche portera sur 3 000 à 5 000 enfants et concerne toutes les familles demandant une place dans n'importe quel établissement collectif géré par la commission d'attribution des territoires engagés dans la recherche. **Sa réussite nécessite la participation active d'une trentaine de territoires disposant de plusieurs centaines de places d'accueil et acceptant d'intégrer le projet pour les commissions d'attribution du printemps 2021.** Cela pour garantir une précision statistique suffisante et la représentation des communes de France dans leur diversité.
Voilà pourquoi nous comptons sur votre mobilisation et espérons que vous rejoindrez cette expérience inédite aux enjeux importants. Vous inscririez ainsi votre territoire dans une démarche d'innovation démontrant votre investissement pour la transparence dans l'attribution des places en EAJE et l'égalité des chances dès les premières années de la vie.

Contact : arthur.heim@cnaf.fr

Cnaf- 32 avenue de la Sibelle - 75685 Paris Cedex 14

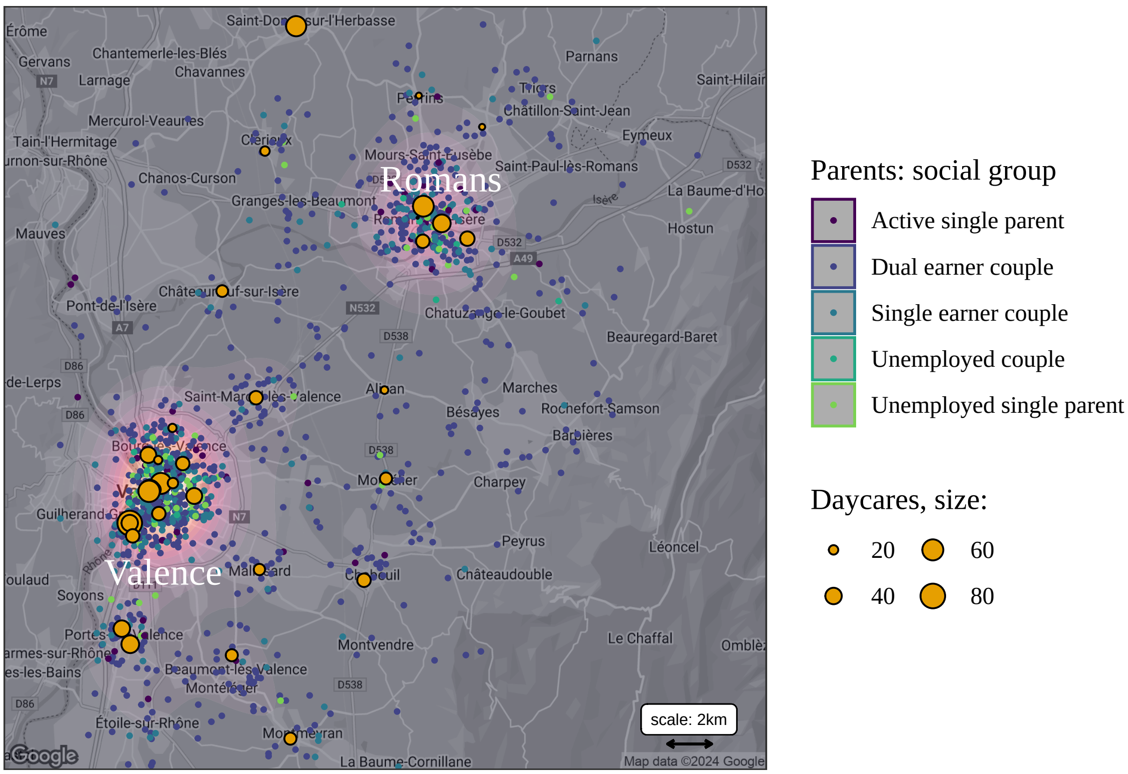
C The market for childcare in Valence Romans

C.I 52 cities in the same daycare markets

Figure C.13: Map of the daycare market in Valence Romans Agglo

Map of the demand, density and daycare's capacities

Demands for the first round in 2023



Sources: ISAJE, Case Study I − 2023, first rounds only.
Notes: Social groups are defined by the local administration with the 'Professional status' criteria. Positions are jittered by .01 degree.
Capacities are those reported on the early−childhood brochure for 2023.

C.II Supply

Table C.3 summarises the aggregated results from various assignment committees. During the four spring committees, daycare centres offered an average equivalent of 415 placements for the first round.

Following the assignment, which could leave vacancies, and considering the decisions of parents accepted in the first round, we initiated a second round with an average equivalent of 124 placements. It's worth noting that most second-round placements predominantly result from family withdrawals rather than vacancies after the procedure.

In 2020, daycare centres supplied an equivalent of 399 placements. However, this first committee occurred amid the initial lock-down, and until August 25th, 2020, the full reopening of creches was uncertain. In 2021, there were nearly 20 fewer placements than in 2020 but more second-round placements. This indicates a higher number of refusals for placements offered in the first round. In 2022, the number of available placements increased significantly. Simultaneously, we will explore how demand also rose substantially that year. Lastly, in 2023, 429 placements were offered.

Table C.3: Available Slots, Number of Sections, and daycares in Different assignment Committees

Year	N daycares		N Buckets		N places	
	1st Round	*2nd Round*	*1st Round*	*2nd Round*	*1st Round*	*2nd Round*
2020	33	29	95	63	399	113
2021	30	31	67	64	380	140
2022	35	34	98	70	451	115
2023	35	32	107	77	429	128
Average					415	124

Sources: ISAJE, Case Study I - 2020 : 2023.
Notes: In 2021, in the first round, childminders were assigned to family daycare centres. For other committees, the early childhood department assigned childminders in advance for families requesting corresponding family daycare centres.

The administration offers guidance in a leaflet and their website[76] allows to map various childcare arrangements. Figure C.13 displays all daycare centres and family childcare in the 54 cities participating in the daycare assignment mechanism. Most childcare facilities are located in the two largest cities Valence and Romans-Sur-Isere.

During the first three years, the number of slots offered to children in the first buckets increased, but there were fewer slots for older children. This situation reversed in 2023.

[76] https://geo.valenceromansagglo.fr

C.III The demand

In each year of the experiment, we processed an average of 1144 applications in the first rounds and 834 in the second rounds. Between 2020 and 2021, the demand increased by 5%, but it was in 2022 that the progression was the strongest: +17% compared to 2021. In 2023, there are 25% more requests than in the 2020 committee.

Table C.4: Attendance at committees from 2020 to 2023

CAMAEn	**Numbers**		
	Total Requests	**Children Presented**	**Families**
April 2020	1,014	1,014	920
June 2020	694	694	642
April 2021	1,064	1,064	991
June 2021	810	810	769
April 2022	1,233	1,233	1,140
June 2022	911	911	860
April 2023	1,267	1,267	1,177
May 2023	919	919	867
TOTAL 1th Round	**4,578**	**4,257**	**3,347**
TOTAL 2th Round	**3,334**	**3,166**	**2,649**
TOTAL unique	**7,912**	**4,681**	**3,551**

Sources: ISAJE, Case Study I - 2020 : 2023 - first rounds only.
Notes:
- The total number of requests counts the registered files, i.e., the number of children.
- Family count identifies the number of unique families at each committee.
- The totals identify the number of unique children and families present.

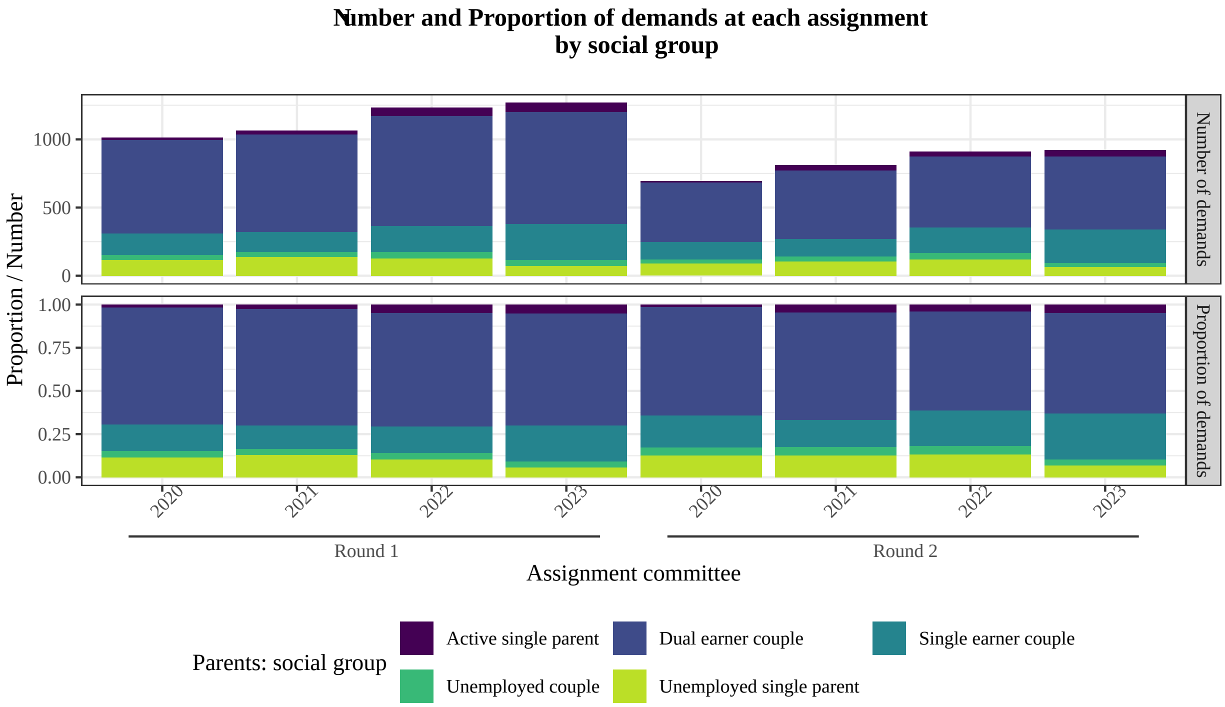
A) The composition of the demand by social group

Table C.5: Composition of the demand by social group across years

	Assignment							
	2020		**2021**		**2022**		**2023**	
	Round 1	*Round 2*	*Round 1*	*Round 2*	*Round 1*	*Round 2*	*Round 1*	*Round 2*
Active single parent	19 (2 %)	11 (2 %)	30 (3 %)	37 (5 %)	61 (5 %)	37 (4 %)	66 (5 %)	45 (5 %)
Dual earner couple	686 (68 %)	436 (63 %)	715 (67 %)	505 (62 %)	809 (66 %)	523 (57 %)	821 (65 %)	536 (58 %)
Single earner couple	156 (15 %)	128 (18 %)	145 (14 %)	126 (16 %)	191 (15 %)	185 (20 %)	266 (21 %)	244 (27 %)
Unemployed couple	36 (4 %)	31 (4 %)	36 (3 %)	39 (5 %)	46 (4 %)	47 (5 %)	41 (3 %)	31 (3 %)
Unemployed single parent	117 (12 %)	88 (13 %)	138 (13 %)	103 (13 %)	126 (10 %)	119 (13 %)	73 (6 %)	63 (7 %)
TOTAL	**1014**	**694**	**1064**	**810**	**1233**	**911**	**1267**	**919**

Sources: ISAJE, Case Study I - 2020 : 2023.
Number and proportion of applicants from each social group in each demand set.

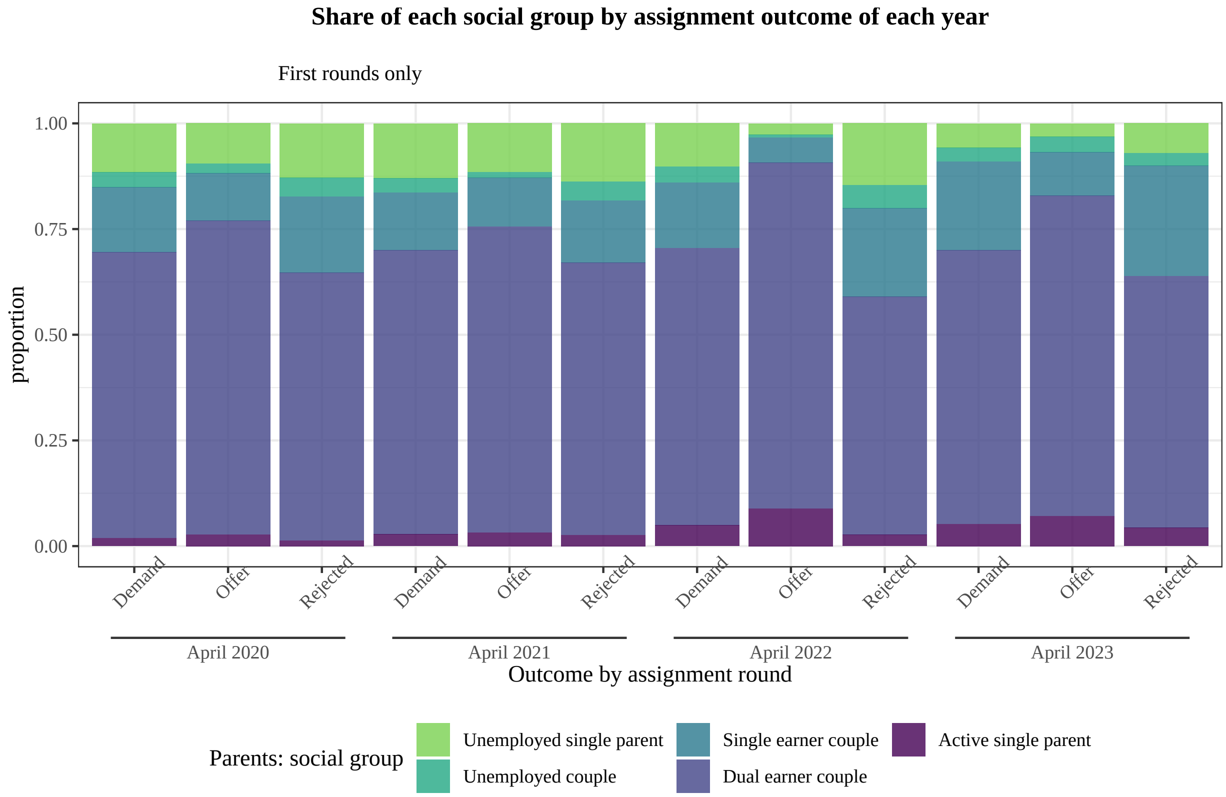
Figure C.14: Demand on the rise



Sources: ISAJE, Case Study I − 2020 : 2023.
Notes:
− Top panel: The number of demands by social group.
− Bottom panel: The proportion of demands at each assignment.

B) Cumulative share of early registration

Figure C.15: Variation of the social composition of assigned and waitlisted groups over the years



Sources: ISAJE, Case Study I − 2020 : 2023 − first rounds only.

C.IV Definition of priorities

A) Criteria

i) Time since registration The committee grants 1/2 point per month since registration. For the research, we needed to adjust these priorities to have enough randomised families. We negotiated with the municipality to retain the criterion but to no longer add 1/2 point per month, instead granting 1.5 points every 3 months. This way, we maintained the general *philosophy* of rewarding families who plan ahead and those who have been waiting for a long time. We just made these benefits less gradual. In 2023, the score uses half a point per month again.

ii) Residence: The committee grants a 4-point bonus for residents of the agglomeration and 0 otherwise. Note that the agglomeration accepts demands from individuals living outside, which is not the case in some territories participating in ISAJE. However, there are almost no family from outside the agglomeration. From 2021 onwards, this weight has been raised to 6 points, requiring parents from outside the agglomeration to wait a year before having the same priority level as residents. What this results in is that parents from outside the agglomeration are *de facto* banned.

iii) Family situation (social or health-related) The *family situation* is taken into account in three variables, and the weights can therefore be added to the others.

- Criteria related to the **parents' situation** can provide up to 4 points, awarded to unemployed parents actively searching for a job (2 points), undergoing job relocation or moving in the territory (1 to 2 points), facing long-term illness, disability, or social support (2 to 4 points).
- Criteria related to the **children's situation** pertain to their health and disability (2 to 4 points) or to the family receiving children related social support (1 to 4 points).
- Criteria related to "**siblings**" grant a 4-point bonus for demands for family reunification in daycare centres, *i.e.*, for families with a child already enrolled in a facility, and an additional 2 points for families with multiple demands (*e.g.* simultaneous demands for twins or siblings).

iv) Social groups: family structure and employment The criteria for *the family's professional situation* combine *activity* and *household structure* to divide demands into 5 mutually exclusive categories:

1) Single-parent in employment receive 4 points, the maximum.
2) Dual earner couples receive 3 points.
3) Single earner couples receive 2 points.
4) Single-parent without employment receive 2 points.
5) Couples with unemployed parents receive 1 point, the minimum.

This criteria defines what we call *social groups*. Table C.5 in Appendix A reports the number and share of each social group every years.

For the following years, the early childhood department revised the weights for some criteria. From 2022 onwards, the priorities for active parents have been dramatically increased. Weights for dual-earner couples have been multiplied by 2.6 going from 3 to 8 points, those of active single-parents by 2.25 going from 4 to 9. Inactive single-parent families and couples now receive 2 or 3 points, while unemployed couples still receive 1 point.

B) Distribution of other priority criteria across groups

Table C.6: Distribution of priority criteria across social groups over each year

	Social groups	Share						N
		children priority	*parents priority*	*multiple demands priority*	*sibling group priority*	*Resident priority*	*early registration (>9 months)*	
2020	*Active single parent*	4.5	36.4	0.0	0.0	100.0	13.6	**22**
	Dual earner couple	2.1	5.1	3.2	3.1	99.5	24.5	**746**
	Single earner couple	12.0	21.0	0.0	1.8	100.0	16.2	**167**
	Unemployed couple	23.7	57.9	0.0	0.0	100.0	13.2	**38**
	Unemployed single parent	4.0	21.6	8.0	3.2	99.2	17.6	**125**
2021	*Active single parent*	6.8	15.9	0.0	4.5	100.0	0.0	**44**
	Dual earner couple	3.0	5.2	4.0	4.8	99.7	12.0	**724**
	Single earner couple	6.4	19.2	1.3	3.8	100.0	5.8	**156**
	Unemployed couple	13.3	37.8	11.1	2.2	100.0	4.4	**45**
	Unemployed single parent	9.1	35.0	7.0	2.1	95.1	9.1	**143**
	Active single parent	0.0	18.2	15.2	0.0	100.0	13.6	**66**
	Dual earner couple	0.5	3.3	9.0	3.1	98.4	14.8	**815**
	Single earner couple	6.0	25.1	14.1	1.0	100.0	11.1	**199**

Sources: ISAJE, Case Study I - 2020 : 2023.
Notes : We identify files at the children level to avoid counting repeated demands for the same child. Totals indicate the number of unique application per children over all years.

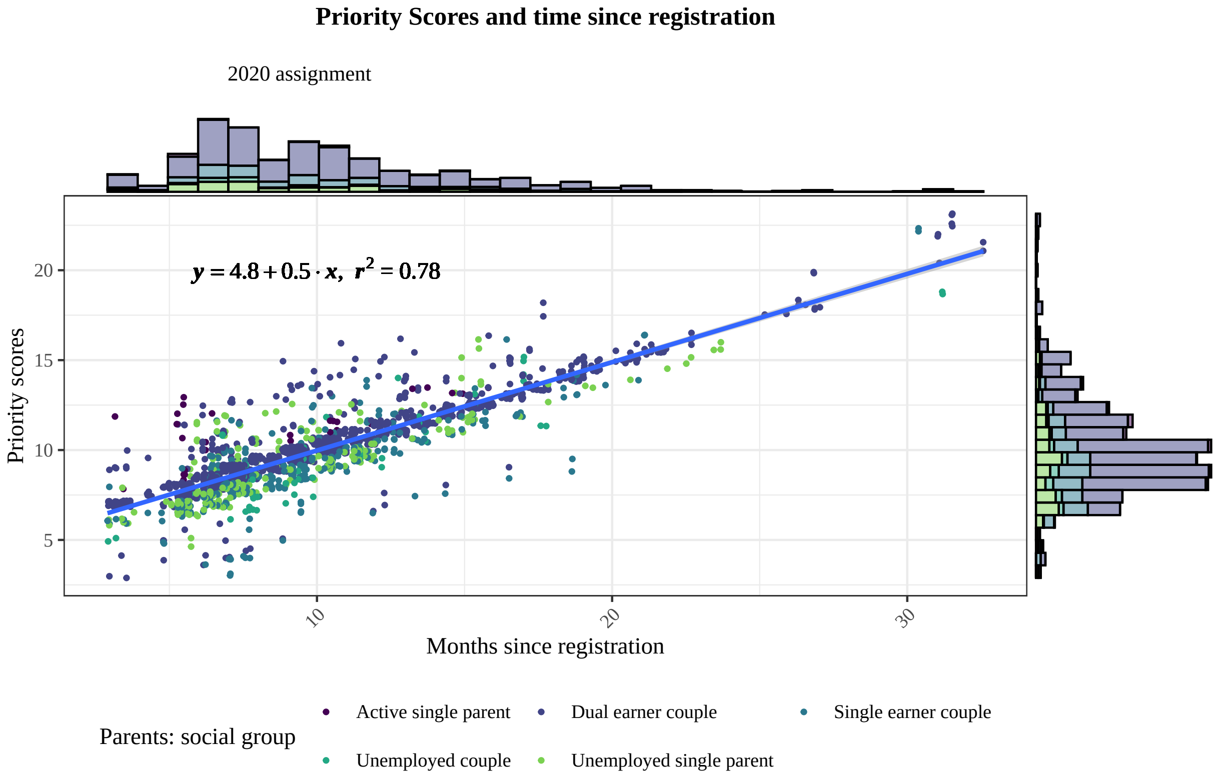
Table C.6: Distribution of priority criteria across social groups over each year

	Social groups	**Share**						**N**
		children priority	*parents priority*	*multiple demands priority*	*sibling group priority*	*Resident priority*	*early registration (>9 months)*	
2022	*Unemployed couple*	4.9	41.5	7.3	0.0	97.6	7.3	**41**
	Unemployed single parent	11.7	39.1	14.1	1.6	99.2	7.0	**128**
2023	*Active single parent*	9.1	19.7	6.1	0.0	100.0	7.6	**66**
	Dual earner couple	2.0	4.8	6.4	2.8	98.9	17.3	**784**
	Single earner couple	16.8	29.0	13.4	3.1	100.0	21.0	**262**
	Unemployed couple	53.5	72.1	27.9	0.0	100.0	25.6	**43**
	Unemployed single parent	38.8	61.2	20.9	0.0	100.0	10.4	**67**

Sources: ISAJE, Case Study I - 2020 : 2023.
Notes : We identify files at the children level to avoid counting repeated demands for the same child. Totals indicate the number of unique application per children over all years.

C) Variation of priorities with time since registration in 2020

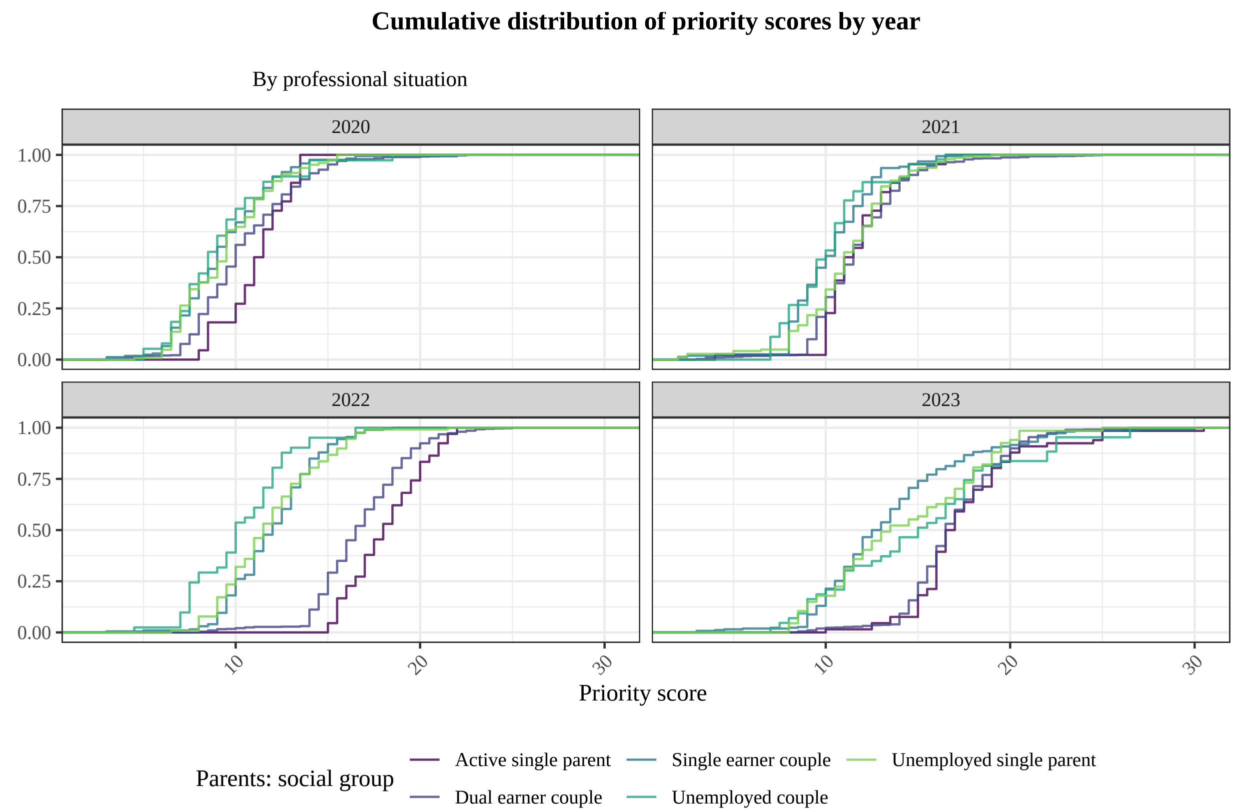
Figure C.16: Variation of priority scores with time since registration in 2020



Sources: ISAJE, Case Study I − 2020, first round only.
Notes: We present the scatterplot of the priority score against the number of months since registration at the time of the committee. The line is the prediction of the linear regression associated with the scatterplot. The colors represent social groups. We add the marginal distributions of both variables at the outset, color coded by social groups.

D) Evolution of the cumulative distribution of priorities over the years

Figure C.17: Structural changes in the composition of demand and priorities



Sources: ISAJE, Case Study I − 2020 : 2023 − first rounds only.

C.V Preferences

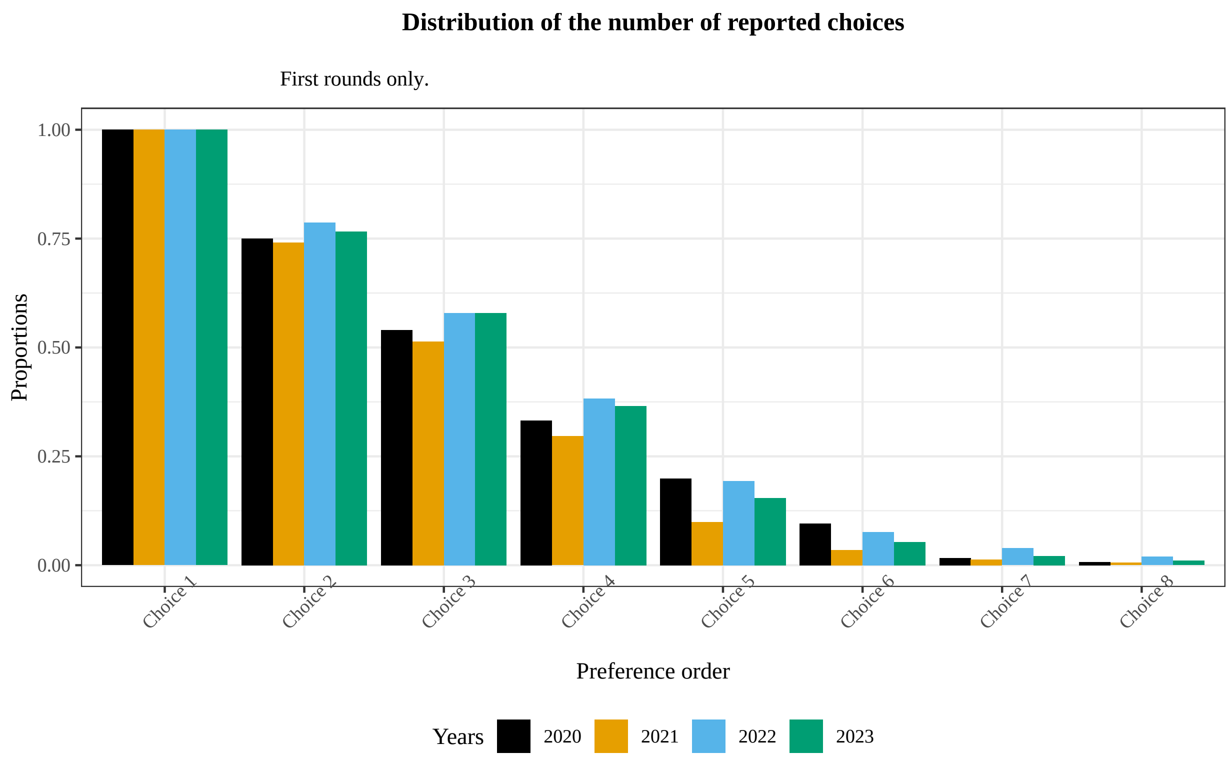
Parents can rank up to 8 wishes and figure C.18 shows their distribution. Most parents report one or two choices. This means that parents' choices are not limited by the 8 preferences, and families find few options acceptable. The committee tries to allocate families from their most preferred to least preferred options, and those who provide fewer preferences have fewer opportunities to be qualified. It all depends on how strong the competition is in the age groups of these childcare facilities.

Finally, we need to analyse the requested days by families. For this, we estimate the average proportion of requests for each day in the first rounds of the committees by professional situation. Figure C.19 presents these results in separate panels for each year.

In 2020 and 2021, aside from differences in average proportions between groups, the patterns of requested days in the week are similar between groups, with only Wednesday being consistently less requested. Starting in 2022, the requested days evolve for some groups: in 2021, unemployed couples and single-parent families mainly request days at the beginning of the week and fewer towards the end of the week. The following year, we see similar requests to 2020 and 2021.

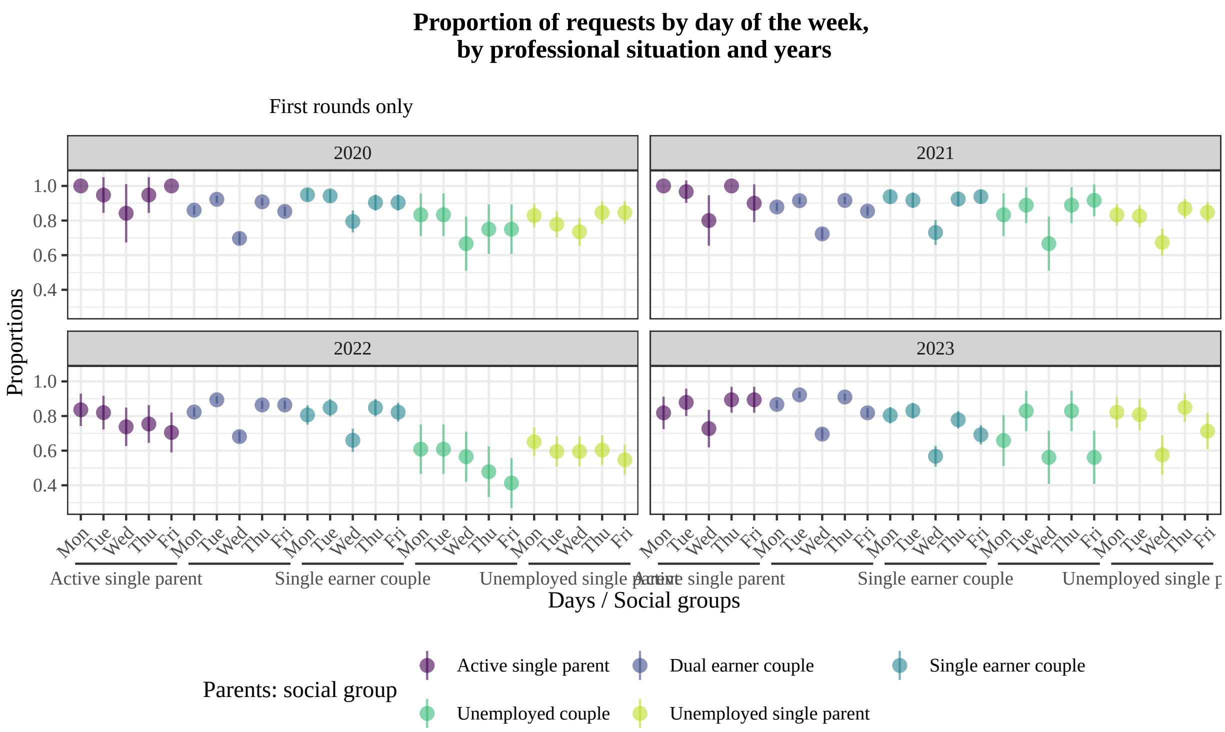
For all families, Wednesday remains the least requested day, and working families request more days than other families.

Figure C.18: 50% of families report 3 choices or less



Sources: ISAJE, Case Study I − 2020 : 2023 − first rounds only.
Notes: In 2020, 75% of applications ranked at least 2 daycares

Figure C.19: Changing patterns in the demands over weekdays over years and social groups

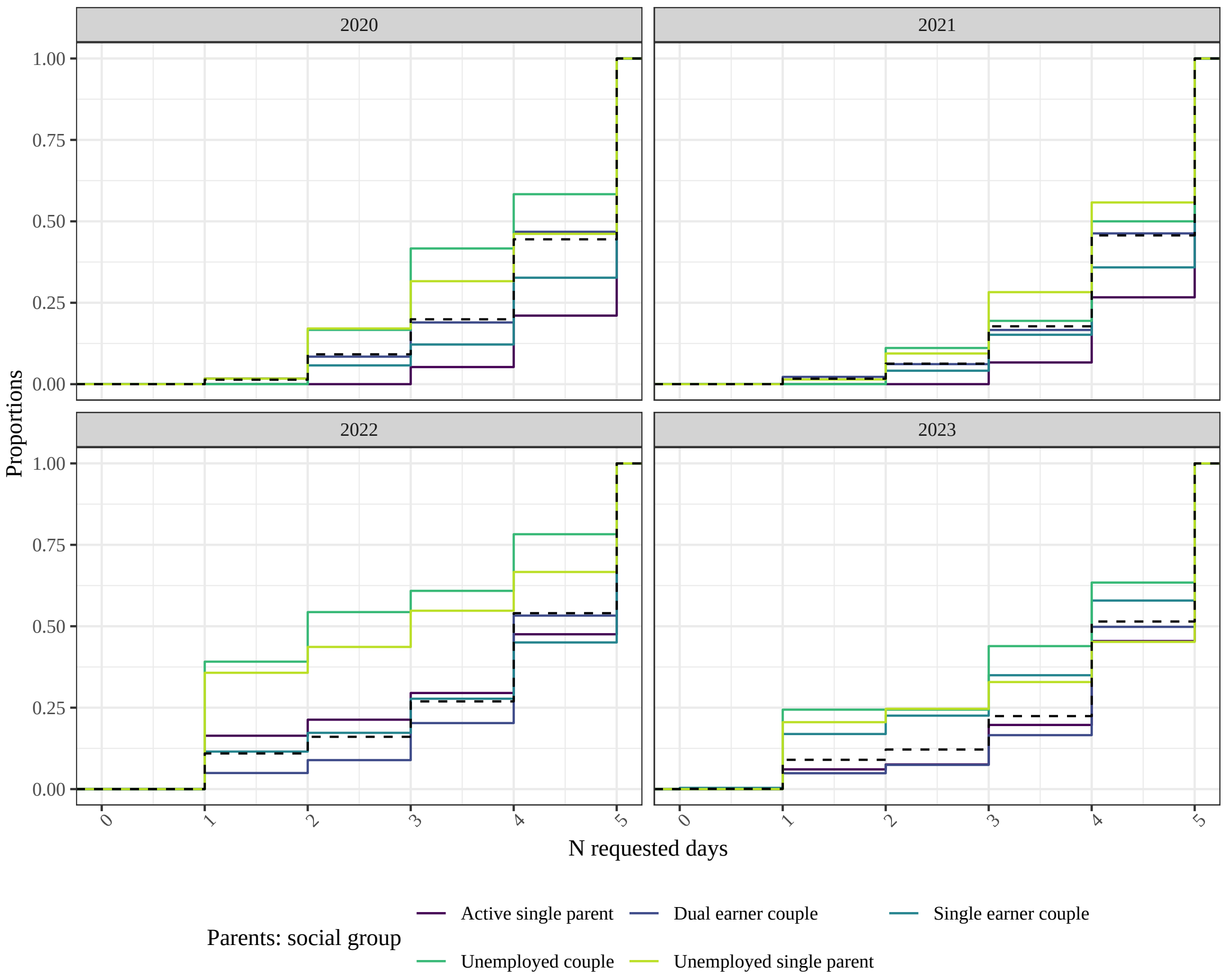


Sources: ISAJE, Case Study I − 2020 : 2023 − first rounds only.
Notes: − Points indicate the average proportion of requests by day, committee year, and social groups. Error bars represent the 95% confidence interval.

Figure C.20: 2022 : higher demands for few days among those without a jobs

Cumulative distribution of the number of requested days by social groups

First rounds only



Sources: ISAJE, Case Study I − 2020 : 2023 − first rounds only.
Notes: empirical cumulative distribution of the number of requested days by social group.
Dashed line represents the ECDF over all demands.

A) Parents reported preferences in a *wide* marketplace

Figure C.21 plots the *heat map* of the average number of applications over years for each daycare and preference order rank. This visualization shows the 5 largest daycare centres are also the 5 most demanded across all ranked choices. Most facilities are located in the two main cities (Valence and Romans-Sur-Isère) and distance to other daycare centres may be large. In the map presented in Figure C.22, we show the flows of reported choices with proportional width by $\log^{10}$ flow size, mapped from parents' home towns to each daycare centres, colour-coded by rank. The two urban centres serve as major hubs, concentrating nine of the most sought-after and, consequently, most congested daycare centres. Parents out of the two large urban centres tend to choose the closest daycare as first choice and then others in the main urban centres. In Figure C.23, we present the distribution of straight-line distances to various types of childcare centres. Registered parents live close to the most demanded daycare centres, especially the most congested ones. This is endogenous selection at play, and one should not infer that distance deter parents. Only that conditional on registration, parents live very close to their chosen daycares. It is likely that parents moved-in partly because of nearby childcare infrastructures (Lawrence, Root, and Mollborn 2015). Residential mobility is influenced by amenities for children, childcare and school in particular[77].

Parents do not seem constrained by the number of options. Many could report more choices but choose not to. So which daycare centres do they chose ?

Figure C.21 plots the *heat map* of the average number of applications over years for each daycare and preference order rank. We rank them from the most demanded – across all choices – to the least. We also add the publicised childcare capacity to the left and the total demand by capacity. This visualization shows the "*most popular*" daycare centres, their capacity and at what rank they are chosen. In this marketplace, the 5 largest daycare centres are also the 5 most demanded across all ranked choices. However, *clé des champs* is much more reported as a first choice and far less as other choices. *Pablo Neruda* is the highest second choice.

To better understand these patterns, we first proceed with a geographical analysis. Recall that this market pools 54 cities over 940 km^2. Most facilities are located in the two main cities (Valence and Romans-Sur-Isère) and distance to other daycare centres may be large. In the map presented in Figure C.22, we show the flows of reported choices with proportional width by $\log^{10}$ flow size, mapped from parents' home towns to each daycare centres, colour-coded by rank. We also display some note-worthy establishments : the top 5 most demanded and the top 5 most congested daycare centres.

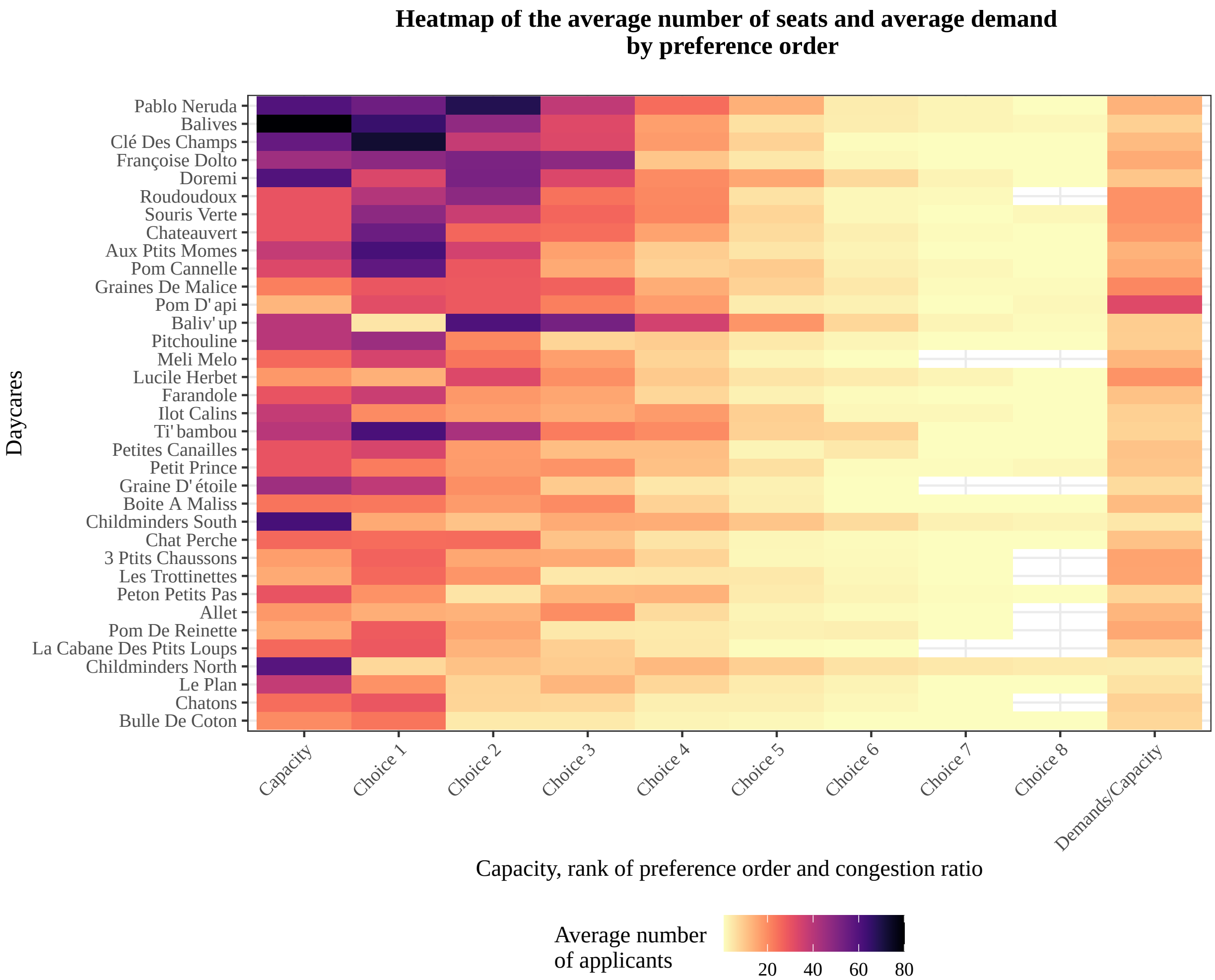
This map reveals three compelling phenomena. First, it's clear that the two urban centres serve as major hubs, concentrating nine of the most sought-after and, consequently, most congested daycare centres. Second, despite being the largest, these daycare centres are often ranked as 3rd or higher choices by parents residing away from the urban centres. Parents out of the two large urban centres tend to choose the closest daycare as first choice and then others in the main urban centres. Third, the third most demanded childcare centre, *Clé des Champs*, located to the north, predominantly attracts first-choice demands from parents in Romans-Sur-Isère. This highlights the significance of the geographic distribution of daycare centres in shaping demand and the strategies parents employ. However, this map does not fully capture the dynamics, hiding those of parents residing within the urban centres.

In Figure C.23, we present the distribution of straight-line distances to various types of childcare centres. The top panels display the distribution for the five most demanded and five most congested daycare centres. In the lower-left panel, we depict the distance distributions for parents who request "crèches familiales," which consist of childminders appointed by the administration. While these childcare facilities are typically located in the children-parents area ("Lieu Parent Enfant"), our focus is primarily on the distribution of choices within this context. As already observed, demand for childminders mostly comes from low rank choices reflecting parents preferences for collective childcare (Cartier et al. 2017). The final panel illustrates the distance distribution for other childcare options.

The distribution of the demand for top 5 most demanded daycares is bimodal only because of *clé des champs* but otherwise, all distribution shows that registered parents live close to the most demanded daycare centres. That is even more striking among the most congested childcare centres.

[77] Avery and Pathak (2021) model the externality of introducing a school choice market on the housing market and describe complex equilibria. In France, Fack and Grenet (2010b) showed that the quality of public schools in Paris increases house prices around.

Figure C.21: Preferences, capacity and congestion

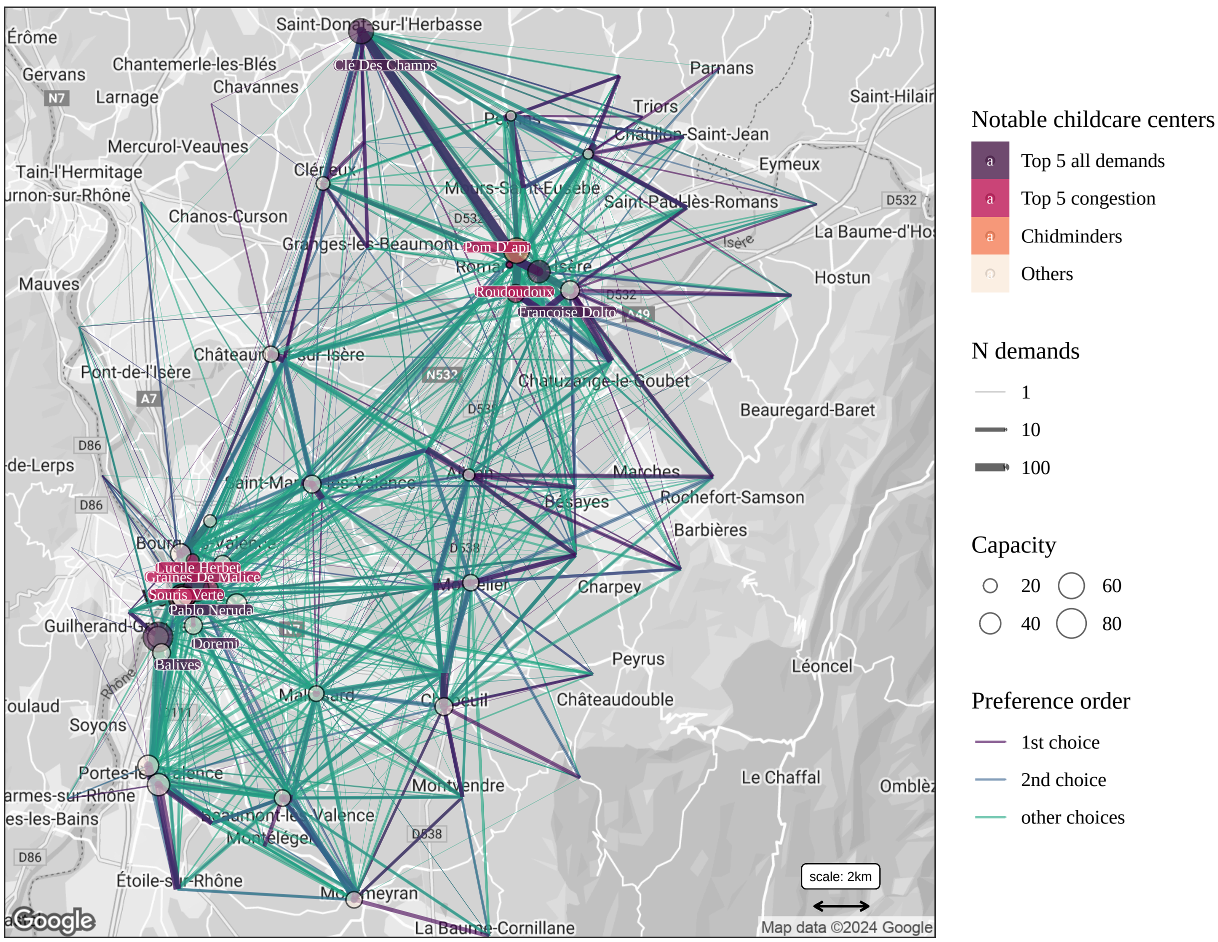


Sources: ISAJE, Case Study I − 2020 : 2023 − first rounds only.
Notes:
− Capacity as reported on the early−childcare page of the adminstration website.
− Average number of applications submitted each year to each daycare by preference order.
− Ratio of the mean total demands over theoretical capacity.

Figure C.22: Demand flows from different cities by preference order

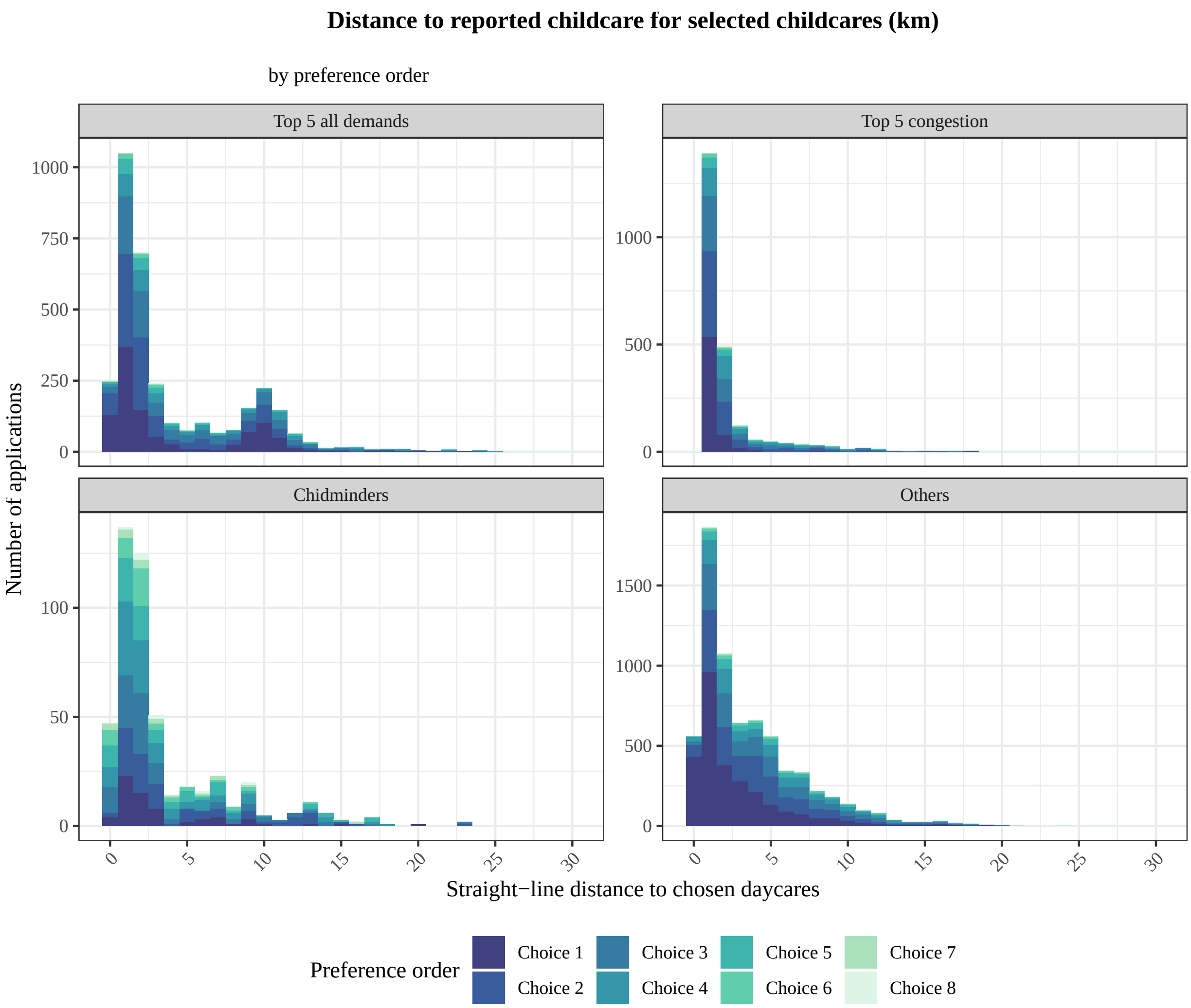
Flow map of demands to ranked daycares by preference order

Demands for the first rounds



Sources: ISAJE, Case Study I − 2020 : 2023 − first rounds only.
Capacities are those reported on the early−childhood brochure for 2023.
All demands are grouped by their home town. Flows are proportional to the log10 of the number of files from the same place.

Figure C.23: Distribution of the distance from home to reported daycare centres by rank and for different group of childcares



sources: Sources: ISAJE, Case Study I − 2020 : 2023 − first rounds only.
Note: Distance from home to reported daycares by group of daycares and rank in the preference order.
Childminders designates 'crèches familiales' and their distance is less meaningful.

D Descriptive statistics on the assignments

D.I Table of aggregate results

Table D.7: Results of automated assignment in VRA over the four years of experiment

		Total					Proportions			
Year	**Round**	Number of demands	Maximum open days	Minimum open days	Average offered seats	Number of assigned families	% Offer	% 1st choice	% 2nd choice	% 3rd choice
2020	1	1,014	440	344	392	400	39.4	64.0	17.2	11.0
	2	694	162	80	111	98	14.1	58.2	18.4	12.2
2021	1	1,064	427	333	380	373	35.1	67.6	18.0	9.4
	2	810	212	97	140	108	13.3	50.9	18.5	15.7
2022	1	1,233	509	401	442	450	36.5	60.9	24.2	10.2
	2	911	175	75	109	109	12.0	52.3	16.5	13.8
2023	1	1,267	488	382	429	408	32.2	58.8	20.8	12.0
	2	919	197	91	128	115	12.5	49.6	25.2	11.3

Sources: ISAJE, Case Study I - 2020 : 2023.
The proportions of choices served are calculated among those accepted.
Reading: In 2023, in the first round, the procedure offers a spot to 408 out of 1267 registered applications, which is 32.2% of demands satisfied, among which 58.8% received their first choice.

D.II Comparison of Accepted and Waitlisted Families' Profiles

Table D.8 presents the average values of various participants' characteristics over the 4 years of the experiment, based on their admission status with the SOFM of that year. We keep only one observation per child.

The results presented in this table are informative. First, in general:

- There are significant differences in virtually all included measures. The only ones with no difference concern the number of requested days.
- The characteristics weighted in the priority scores increase the probabilities of admission.

Regarding family strategies:

- Families who submitted only one wish, on average, have a 6-point lower probability of being admitted.
- Families who submitted at least 3 wishes are 10 points more likely to receive an offer than those who ranked less daycares.
- Admitted families enroll much earlier, for a large part before delivery.
  - Admitted children are 1 month older.
  - Children admitted in September have been registered for an average of 13 months, 3.5 months longer than those in the waiting list, on average.
  - The average child age at the time of registration is 2.5 months younger among assigned families than among waitlisted ones.
  - 65% of admitted children were enrolled before birth, compared to 47% of rejected children.

These observations highlight the direct impact of parental anticipation on the outcomes.

Regarding professional situations, dual-working couples and active single-parent families are significantly more represented among the admitted families than among those on the waiting list. However, for other types of families, the differences are reversed.

The last row of the table indicates the share of families whose assignment depends on the outcome of the lottery, which is necessary for the ISAJE research[78]. Overall, the assignment is largely determined by priorities and capacity constraints, as only 20% of admitted families have an assignment dependent on the lottery draw, and only 10% of those on the waiting list. This confirms that the experimental conditions did not have a significant effect on the course of the assignment committees. We did not disrupt much the assignment by introducing lotteries.

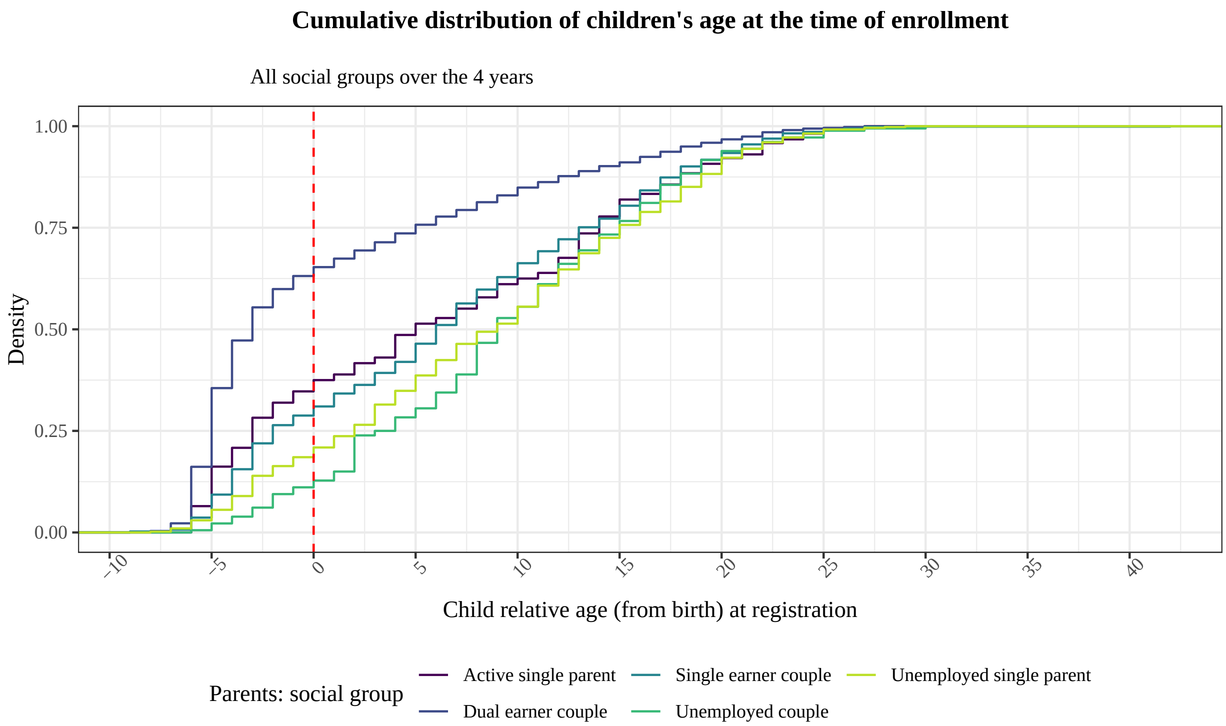
[78] It is obtained by simulating 1,000,000 assignments while fixing the entire problem and changing only the lottery draw. For each record, we then calculate the proportion of these one million simulations in which it is admitted or on the waiting list.

Table D.8: Comparison of assigned and waitlisted families over the 4 years in the experiment

	Offer (N=1462)		Rejected (N=3219)		Diff. in Means	Std. Error
	Mean	*Std. Dev.*	*Mean*	*Std. Dev.*		
Child Age (months)	13.39	7.99	12.21	8.74	-0.99***	0.29
Number of Preferences	3.22	1.66	2.87	1.58	-0.38***	0.06
Share with Only One Preference	0.19	0.39	0.24	0.43	0.06***	0.01
Share with at Least 3 Preferences	0.63	0.48	0.54	0.50	-0.10***	0.02
Average Number of Days Requested	4.16	1.06	4.10	1.22	-0.06	0.04
Share Requesting 5 Days	0.49	0.50	0.53	0.50	0.03*	0.02
Share Requesting at Least 3 Days	0.20	0.40	0.23	0.42	0.02	0.01
Average Months Since Registration	13.06	4.60	8.82	3.71	-3.55***	0.13
Average Child Age at Registration	0.77	8.60	3.80	9.07	2.54***	0.31
Share Registered Before Birth	0.65	0.48	0.47	0.50	-0.16***	0.02
Share Dual earner couple	0.76	0.43	0.61	0.49	-0.15***	0.02
Share Single earner couple	0.09	0.29	0.20	0.40	0.11***	0.01
Share Unemployed couple	0.02	0.13	0.04	0.20	0.03***	0.00
Share Unemployed single parent	0.07	0.25	0.11	0.32	0.05***	0.01
Share Active single parent	0.06	0.23	0.04	0.18	-0.03***	0.01
Share Random	0.20	0.40	0.11	0.31	-0.08***	0.01

* = p<.1, ** = p<.05, *** = p<.01
Sources: ISAJE, Case Study I - 2020 : 2023 - first rounds only.
The reference for all date calculations is September 1st.
Note: We present means and standard deviations in each group and the mean difference and associated standard error, accounting for stratification by assignment round.

Figure D.24: Cumulative empirical distribution of the age of the child at month of registration



Sources: ISAJE, Case Study I − 2020 : 2023.
notes:
− One observation is kept for each child presented to one or the other of the two committees,
− The variable analysed is constructed as the difference, rounded in months,
between the date of birth (actual or expected) and the date of registration.

D.III Wastefulness: Illustration with the First Round of the 2023 committee

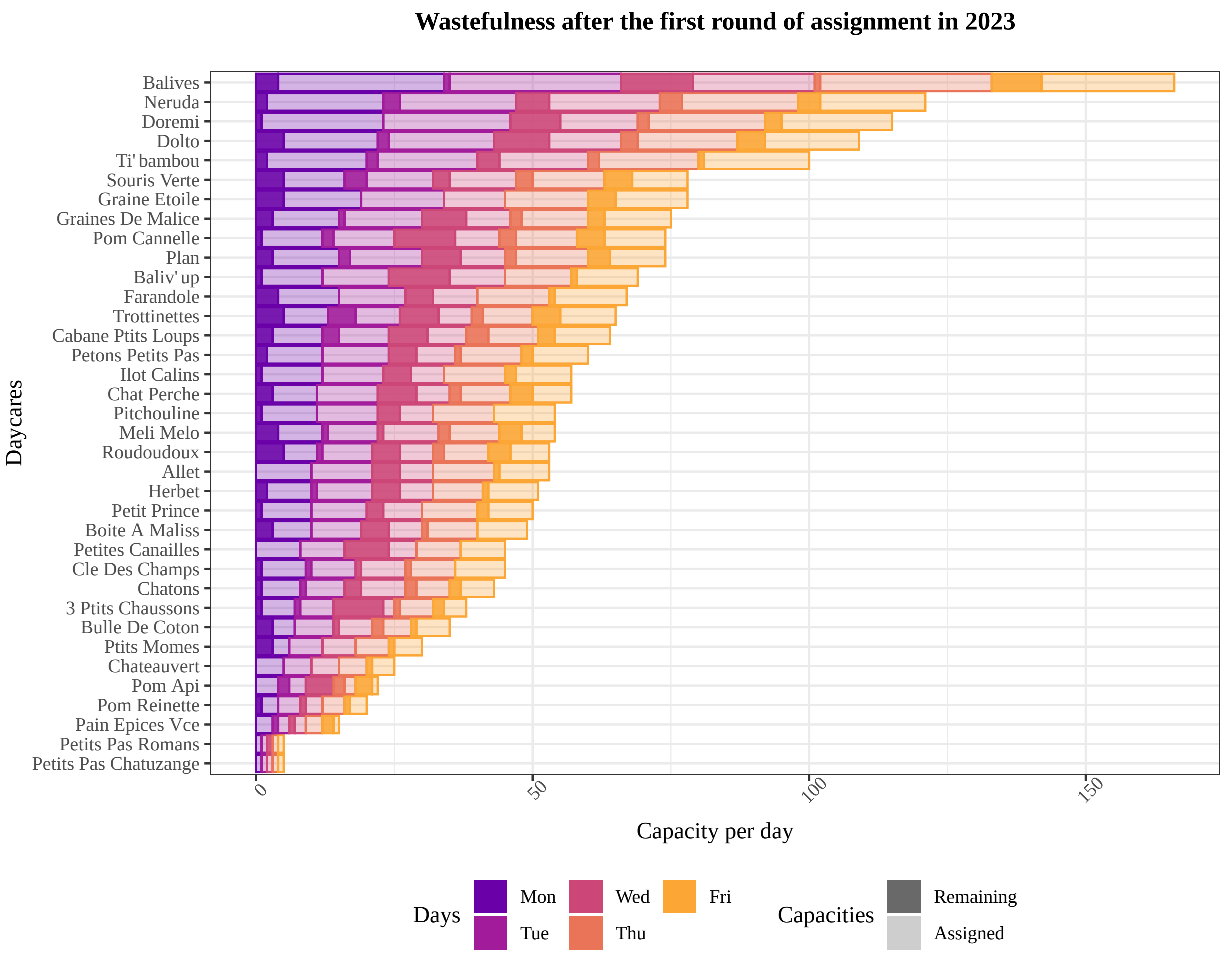
As explained earlier, our algorithms allow applications to be ranked within each bucket based on priority scores, and we continuously check if there are enough capacities in each day to accommodate the received applications. With the CAM algorithm, as soon as a day reaches its capacity threshold, the subsequent applicant that requires that day sets the final cutoff. If we use an alternative definition of *fairness* and allow weak accommodation, more children may still be accepted. However, it's important to note that the majority of demands are for full-time slots, with only a few part-time demands that can be combined (Especially in 2022, see Figure C.20 in the Appendix). Consequently, once the most requested days can no longer accommodate new demands, some buckets are left with available capacities on specific days.

In Figure D.25, we illustrate the wastefulness across daycare centres in the 2023 first round. The majority of capacities are successfully assigned, but there are still some vacant spots in all facilities. Notably, given the lower demand for Wednesdays (Figure C.19 in the Appendix), most remaining slots are on that particular day. During the 2023 round, rejections primarily result from capacity exhaustion on Tuesdays and Thursdays.

These available spots become part of the pool for the second round after families receive notification of the results approved by the assignment committee. This validation process includes confirmation with the early childhood service upon notification. More importantly, it entails scheduling appointments with the directors of the childcare facilities. During these contractual negotiations, family needs can be adjusted, and any days left vacant due to withdrawals and adjustments are reintroduced in the second round.

Essentially, the "*waste*" in the first round finds a second chance for assignment in the second round. Moreover, the spots from the second round are repurposed in September to facilitate new assignments, especially for providing occasional care. This coordination between two "*markets*," each catering to different needs, allows for more efficient resource utilisation. Nonetheless, one might question whether we can achieve even more efficiency by employing alternative methods, such as the KDA procedure.

Figure D.25: Used and remaining capacities after the first round of assignment in 2023



Sources: ISAJE, Case Study I 2023.
Notes: We sum the available slots per day in the sections and subtract those assigned.

D.IV Comparison of the final assignment with results of the algorithm

Table D.9: Deviation from the family optimal fair matching in the final assignment

SOIFFA	Final assignment				
	Same as SOIFFA	*Unmatched*	*Waitlist reassigned*	*Lower ranked choice*	*Higher ranked choice*
1 choice	217	6		14	
2 choice	73	2		1	9
3 choice	41	2			3
4 choice	14	1			4
5 choice	9	1			
6 choice	1				1
7 choice	2				
Waiting List	813	3	47		
TOTAL	**1,170**	**15**	**47**	**15**	**17**

Sources: ISAJE, Case Study I - 2023, first round only.
Notes: This table compares the actual assignment with the proposal from the SOIFFA. The first column indicates how many files received the same offer. In the following columns, we indicate how many families were moved from the initial assignment in the corresonding row.
217 families had their first choice with the algorithm, but 6 should have had their first choice and were placed on the waiting list by the allocation committee. 47 families were placed on the waiting list in the SOIFFA and received a proposal instead.

D.V Assignment probabilities by social groups

Table D.10: Probability of receiving an offer by social group

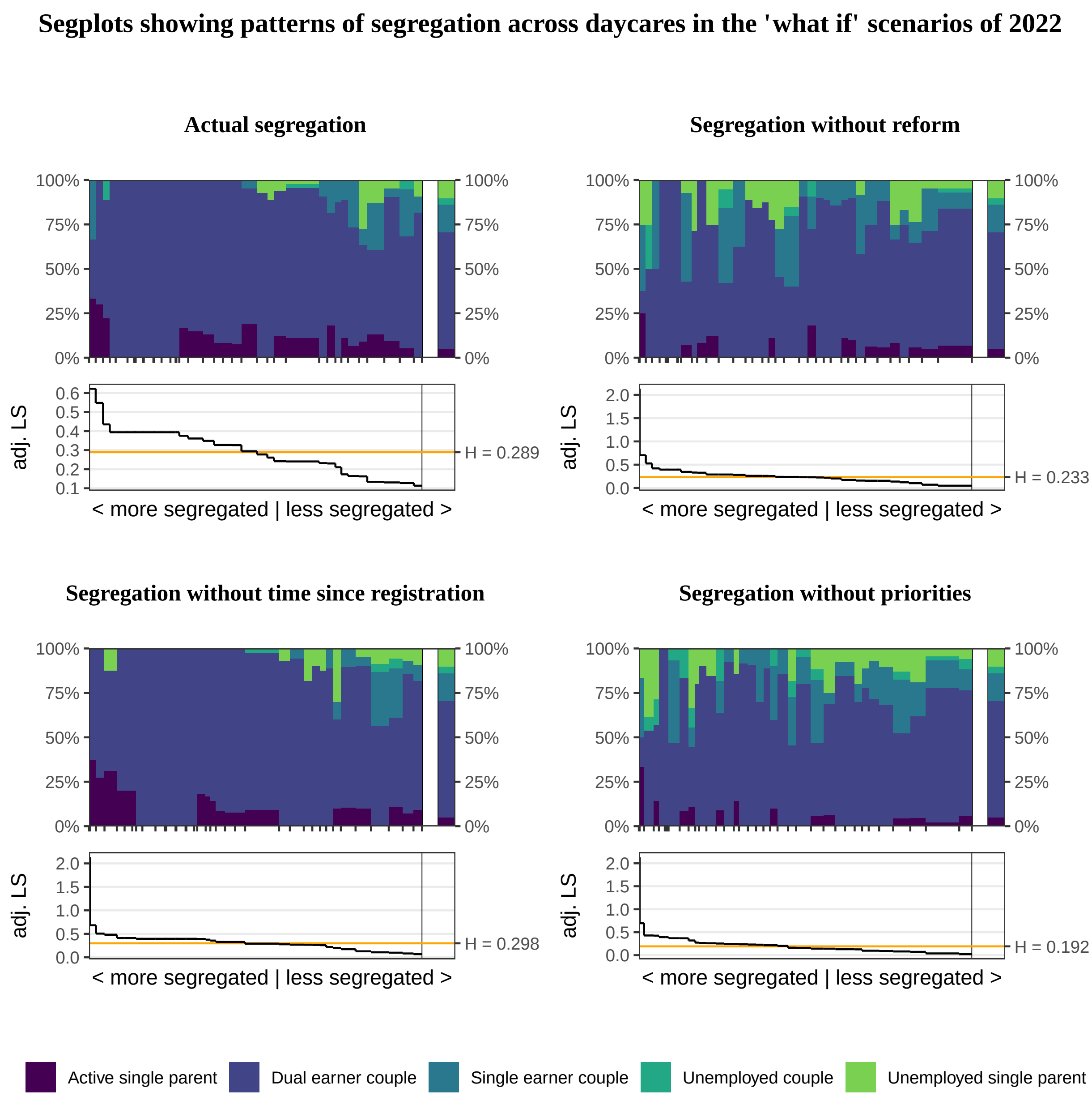
	sample				
	Full sample	*2020*	*2021*	*2022*	*2023*
Active single parent	0.52*** (0.04)	0.58*** (0.11)	0.40*** (0.09)	0.66*** (0.06)	0.44*** (0.06)
	[0.45, 0.60]	[0.36, 0.79]	[0.22, 0.58]	[0.53, 0.78]	[0.32, 0.56]
Dual earner couple	0.41*** (0.01)	0.43*** (0.02)	0.38*** (0.02)	0.45*** (0.02)	0.38*** (0.02)
	[0.39, 0.43]	[0.39, 0.47]	[0.34, 0.41]	[0.42, 0.49]	[0.34, 0.41]
Single earner couple	0.21*** (0.02)	0.29*** (0.04)	0.30*** (0.04)	0.14*** (0.03)	0.16*** (0.03)
	[0.18, 0.24]	[0.21, 0.36]	[0.22, 0.37]	[0.09, 0.20]	[0.11, 0.21]
Unemployed couple	0.20*** (0.04)	0.25*** (0.07)	0.14** (0.07)	0.07 (0.05)	0.37*** (0.09)
	[0.13, 0.27]	[0.11, 0.39]	[0.01, 0.27]	[-0.02, 0.16]	[0.19, 0.55]
Unemployed single parent	0.23*** (0.02)	0.32*** (0.05)	0.31*** (0.04)	0.10*** (0.03)	0.18*** (0.05)
	[0.19, 0.27]	[0.23, 0.42]	[0.23, 0.39]	[0.04, 0.15]	[0.09, 0.27]
Num.Obs.	4578	1014	1064	1233	1267
Std.Errors	by: idfam	by: idfam	by: idfam	by: idfam	by: idfam

Sources: ISAJE, Case Study I - 2020 : 2023 - first rounds only.
notes: OLS estimate without constant of the proportion of families who received an offer using the automatic procedure each year. The coefficients indicate the proportion of families in each category who received an offer, the number in brackets indicates the standard error. In the first round of 2023, 37.6% of parents in a dual-earner couple were offered a place using the automatic procedure.

D.VI Analysis of segregation within daycares

A) Segplots comparing segregation across *what if* scenarii in 2022

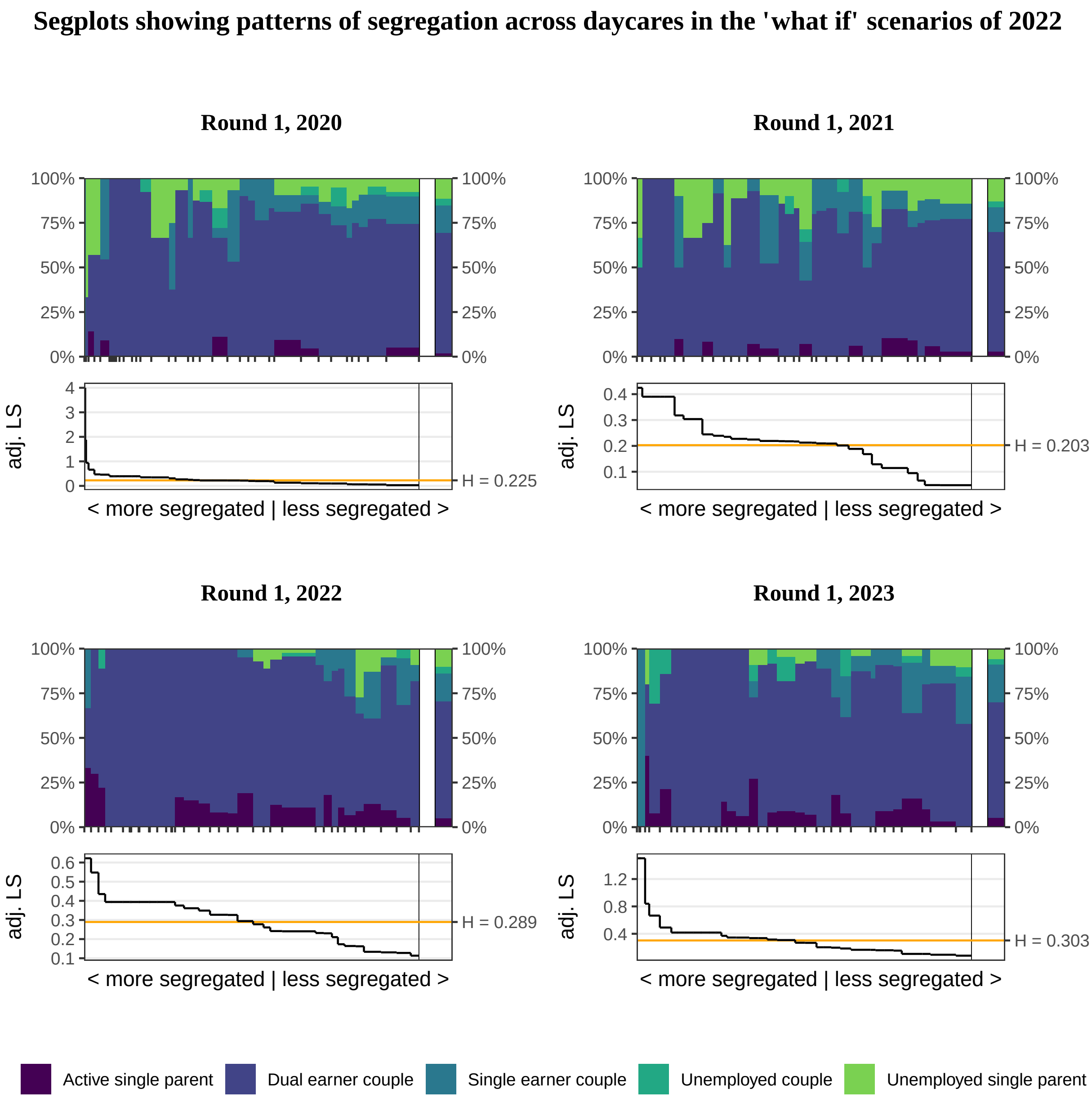
Figure D.26: Segplots across daycares by social groups in the 2022 what if scenarii



Sources: ISAJE, Case Study I − 2022, first rounds only.
Notes: Segplots are defined by Elbers and Gruijters (2024) and implemented in the package 'segregation'.
Each daycare is shown as one individual bar, where the width of the bar is proportional to the size of the daycare.
Each bar shows the distribution of social groups within each daycare.
The rightmost bar depicts the 'reference distribution'. This is the overall group distribution in the demand, i.e., including those in the waiting list. Daycares are ordered in each segplot by decreasing local segregation scores.
The bottom figure shows the adjusted local segregation scores of each unit and the overall H index (horizontal orange line).

B) Segplots comparing segregation through the years

Figure D.27: Segplots across daycares by social groups over the years in the first rounds



Sources: ISAJE, Case Study I − 2020 : 2023 − first rounds only.
Notes: Segplots are defined by Elbers and Gruijters (2024) and implemented in the package 'segregation'.
Each daycare is shown as one individual bar, where the width of the bar is proportional to the size of the daycare.
Each bar shows the distribution of social groups within each daycare.
The rightmost bar depicts the ...reference distribution.... This is the overall group distribution in the demand, i.e., including those in the waiting list. Daycares are ordered in each segplot by decreasing local segregation scores.
The bottom figure shows the adjusted local segregation scores of each unit and the overall H index (horizontal orange line).

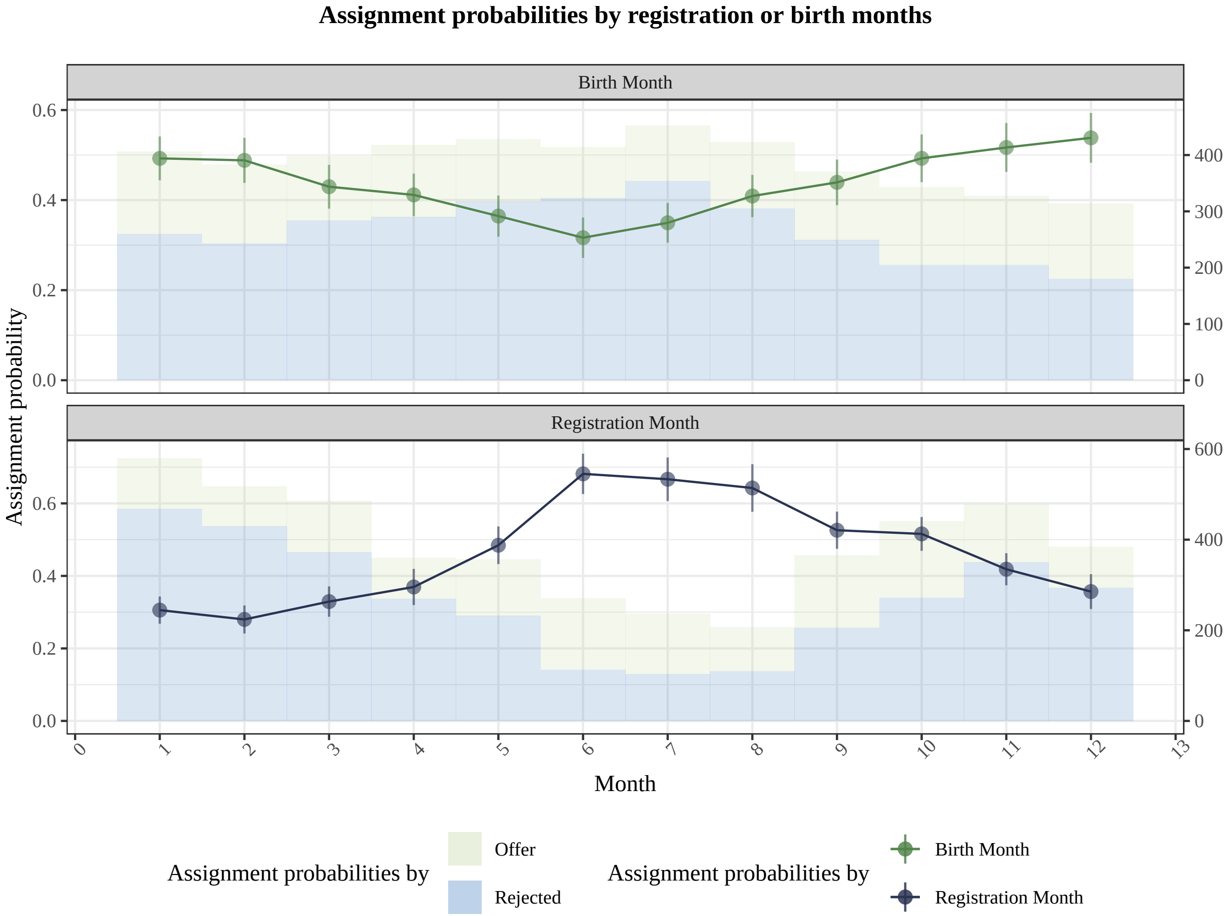
D.VII Assignment probabilities and months of registration and birth

Example 1.4 (The unfair advantage of well timed pregnancy)**.** Take two dual-earner couples living in the urban area, without other particular priorities[79]. A parent of the first couple becomes pregnant in February, discovers the pregnancy in April, and declares it in May. If all goes well, the child should be born in October. Given the school year, this family can expect a slot in September when the child is about 11 months old. Their dominant strategy (if they can) is to register in May. When the assignment occurs in the following April, they have been registered for 11 months, which adds 5.5 points to the 7 they have as dual-earners from the urban area. In 2020, this score puts them in the top 20% of priorities, significantly increasing their chances of obtaining a slot, provided they can wait until the child is 11 months old and do not change their mind by then.

The parent of the second couple becomes pregnant in October, discovers the pregnancy in December, and declares it in January for a child expected in July. By the start of September, this child will be about 3 months old. If this family follows the same strategy by registering as soon as they declare the pregnancy, their score will only be increased by 1.5 points in April, placing them below 55% of the demands. Their chances of obtaining a slot are therefore considerably reduced. However, if they can wait for another year until the child is 15 months old, they will have a higher priority.

[79] that is, the situation of 59.6 % of demands.

Figure D.28: Between 2020 and 2023, the probability of admission is 1/3 for families registering in January and 2/3 for those registering in June.

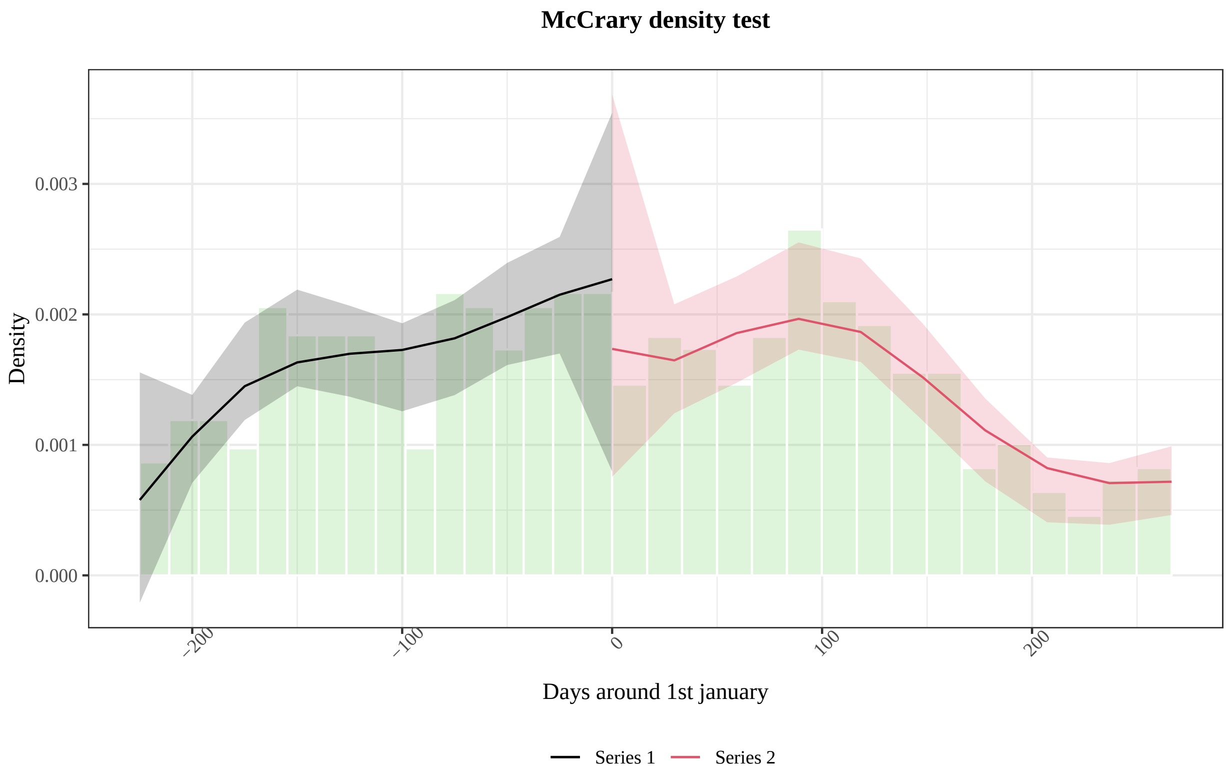


Sources: ISAJE, Case Study I − 2020 : 2023 − first rounds only.
Notes: The points indicate the proportion that received an offer based on
− the month of registration (top panel)
− the child's birth month (bottom panel)
The value of the histograms is readable on the right scale and indicates the number of observations for each month, colour−coded by admission status after the automatic procedure.
Over the period 2020−2023, children registered in June have a probability of admission of nearly 60%, while children born in October have a probability of admission of 40%.

E Additional results on Case study II

E.I McCrary density test of discontinuity

Figure E.29: No discontinuity in the densities of date of birth



Sources: ISAJE, Case Study II.
Notes: Validation test following McCrary (2008) of no discontinuity in the density around the cutoff.
Non parametric estimations following Cattaneo et al. (2014).

F Algorithms, proofs and illustrations of the main theoretical results

F.I Formal presentation of the algorithms

A) The student proposing deferred acceptance algorithm

Intuitively, we *break* daycares into individual buckets (seats) and transform preferences over daycares into preferences over buckets. This allows to define priorities for each buckets easily allowing different ordering of students within a daycare to respect diversity constraints. With SPDA, at any moment in the procedure, each bucket (seat) will keep the student that applies to this seat that has the highest local priority on that bucket. If a bucket only accepts students of group g (hard quota) and no such student applied to this seat, this bucket can end up empty. Since this wastefulness comes from eligibility rule, this matching is still stable since other student are not eligible. Soft quotas reduce the risks of such cases but again, this is a policymaker's decision.

Algorithm 1 Student Proposing Deferred Acceptance (SPDA)

Require: $(inputs := (I, S, \mathbf{q}, \theta))$
  $\pi_{is} \leftarrow \rho_{is} + \epsilon_i$ ▷ Priorities with random tie-break
  $k \leftarrow 1$ ▷ Initialisation
  all applicants apply to their first-ranked schools.
  if All schools can accomodate all applicants **then**
    assignment is final
    return $\mu(k)$
  else if at least one school receives more applicants than its capacity **then**
    keep the best applicants in all schools according to their scores π_{is} up to the capacity constraint and permanently reject the others.
    Rejected applicants apply to their next favourite school[a].
    $k \leftarrow k+1$
  end if
  while $k \geq 2$ **do**
    Sort previously held and new applications in each school according to their scores π_{is}
    if the total number of new and previously accepted applicants is greater than its capacity **then**
      keep the best applicants in all schools according to their scores and permanently rejects the others[b]
      applicants rejected by a school in the previous step apply to their next favourite school
      $k \leftarrow k+1$
    else
      Assignment is final
    end if
  end while
  return $\mu(k)$

[a] So their favorite school among those which have not rejected them yet.

[b] Note that the assignment is always provisional, since schools must select the best applicants among both the new ones and the previously accepted ones. Thus, a student accepted at a given step can still be rejected if another student with a better score applies to the school. This is a key element of the algorithm.

B) The cut-off adjustment mechanism

Before describing the algorithm, let us introduce a few notations.

For a school s, let c_s be a **cut-off** of school s which is an integer in $\{0,\dots,N\}$ where $N=|I|$ is the number of applicants. Once ties are broken and the scores π_{is} at school s are obtained, order the applicants by decreasing order of score and let $\tilde{r}_{is}$ be the **rank of applicant** i in this ordering.

Given a collection of cut-offs $c := (c_s)_s$, let:

- $A_i(c) := \{s \in S : \tilde{r}_{is} \leq c_s\}$ be the set of **accessible schools** for applicant i.
- $D_s(s) := \{i \in I : s \in A_i(c)$ and $s \succ_i s'$ for $s' \in A_i(c)\backslash\{s\}\}$ be the **demand set** of school s at the cut-offs c.

In plain words, accessible schools are those where the rank of an applicant i in that school is lower than the cut-off ; the demand set is composed of all applicants in their most favourite accessible school. Cut-offs determine who can apply to a school and, given a collection of cut-offs, applicants apply to their favourite accessible school.

Algorithm 2 Cut-off adjustment mechanism

Require: *Inputs* (I, S, **q**, θ))
  $round \leftarrow 0$ ▷ *Steps 0*
  if Applicant i is ineligible at s **then**
    $\pi_{is} \leftarrow -\infty$ ▷ Or simply negative priorities[a]
  else
    $\pi_{is} \leftarrow \pi_{is} + \epsilon_i$
  end if
  $\tilde{r}_{is} \leftarrow rank(\pi_{is})$ ▷ Define the rank of each applicant in each school
  $c_s(0) = N - \sum_i \mathbb{1}(\pi_{is} < 0)$ ▷ Initialise cut-offs as the number of applicants eligible in that school
  $round \leftarrow 1$ ▷ *steps round* ≥ 1
  while $round \geq 1$ **do**
    $\mu(round) \leftarrow \mu_s(k) := D_s(c(k-1))$ ▷ Demand for school s over the cut-off at round k
    if $\mu(k)$ is feasible **then**
      Exit while
    else
      $\tilde{S}(\mu(k)) := \{s \in S : \exists t, \sum_{i \in \mu_s(k)} d_i^t > q_s^t\}$ ▷ Schools for which $\mu(k)$ is not feasible
    end if
    if $s \in \tilde{S}(\mu(k))$ **then** ▷ Adjust the vector of cut-offs $c(k)$ for schools in $\tilde{S}(\mu(k))$
      $c_s(k) = c_s(k-1) - 1$ ▷ Decrease the cut-off by one rank
    else
      $c_s(k) = c_s(k-1)$ ▷ Leave other cut-offs like they were in the previous step
    end if
    $k \leftarrow k+1$
  end while
  return $\mu(k)$

[a] Another alternative is to just delete that school from the preference list. We discuss such settings more precisely later in the paper.

C) The Knapsack deferred acceptance algorithm

The main description of KDA by their authors is different from this one. We want to emphasise the link between CAM and KDA so we use the cut-off adjustment representation of the deferred acceptance mechanism in the multi-dimensional setting. Kamada and Kojima (2023) show in their appendix that the SOFM can be obtained by this algorithm using their definition of justified envy. The main difference with the CAM algorithm is that we now use **weak-justified envy** to define the **cut-off adjustment function**. In practice, the Matlab functions are based on the definition of KDA from Delacrétaz, Kominers, and Teytelboym (2023) which is empirically *faster* than this version.

Algorithm 3 Knapsack Deferred Acceptance

Require: *Inputs* (I, S, **q**, θ))
  $round \leftarrow 0$ ▷ *Steps 0*
  if Applicant i is ineligible at s **then**
    $\pi_{is} \leftarrow -\infty$ ▷ Or simply negative priorities[a]
  else
    $\pi_{is} \leftarrow \pi_{is} + \epsilon_i$
  end if
  $\tilde{r}_{is} \leftarrow rank(\pi_{is})$ ▷ Define the rank of each applicant in each school
  $c_s(0) = N - \sum_i \mathbb{1}(\pi_{is} < 0)$ ▷ Initialise cut-offs as the number of applicants eligible in that school
  $round \leftarrow 1$ ▷ *steps round* ≥ 1
  while $round \geq 1$ **do**
    $\mu(round) \leftarrow \mu_s(k) := D_s(c(k-1))$ ▷ Demand for school s over the cut-off at round k
    if $\mu(k)$ is **weakly feasible then**
      Exit while
    else
      $\tilde{S}(\mu(k)) \leftarrow \{ s \in S : \exists t \ s.t. \ d_i(t) = 0 \mid d_i(t) + \sum_{i \in \mu_s(k)} d_i(t) > q_s^t \}$ ▷ Schools for which $\mu(k)$ is not weakly feasible
    end if
    if $s \in \tilde{S}(\mu(k))$ **then** ▷ Adjust the vector of cut-offs $c(k)$ for schools in $\tilde{S}(\mu(k))$
      $c_s(k) = c_s(k-1) - 1$ ▷ Decrease the cut-off by one rank
    else
      $c_s(k) = c_s(k-1)$ ▷ Leave other cut-offs like they were in the previous step
    end if
    $k \leftarrow k + 1$
  end while
  return $\mu(k)$

[a] Another alternative is to just delete that school from the preference list. We discuss such settings more precisely later in the paper.

D) SPDA with diversity requirements

Algorithm 4 SPDA with soft quotas

Require: *Inputs* $P = (I, S, \mathbf{q}, \underline{\mathbf{q}}, \theta)$; *Function* (*SPDA*)

procedure Break the problem in soft quotas(P)

For each school $s \in S$ and group g, create $\underline{q}_s^g$ buckets with $c_b = 1$. Let B_s^g the be the set buckets in school s corresponding to group g. Let $B := \bigcup_{s,g} B_s^g$.

Fix an ordering $\gg_s$ for each school over its buckets.

For a bucket $b \in B_s^g$, let $\succ_b$ s.t. $g \succ_b \emptyset \succ_b g'$ for $g' \neq g$.

Define $\tilde{\theta}$ such that :

i) $\tilde{\succ}_i$ is the preferences of i over B ▷ apply definition 1.4 on reported preferences to define preferences over buckets

ii) $\tilde{\pi}_{ib}$ is the score of i in bucket b ▷ apply definition 1.3 on priorities to define priorities within buckets

Let $\tilde{P} := (I, B, c, \tilde{\theta})$ be the assignment problem over buckets.

end procedure

Run SPDA on $\tilde{P}$

Output $\tilde{\mu}$ which is a matching of students to buckets

Let μ be the matching of students to schools s.t. $\mu_i = s_{\tilde{\mu}_i}$ ▷ Map matched students in each buckets to the schools and buckets

return μ

F.II Example and proof of theorem 1

Example 1.5 (Motivating example)**.** Consider a very simple market with T=2 (e.g. half-time/full-time), 2 schools A and B with capacities $q_A = (1,0)$ and $q_B = (1,1)$, and three students i_1, i_2 and i_3 of the same age group with types θ summarised:

$$\begin{aligned} A \succ_{i_1} B\,;\, d_{i_1} = (1,0) \quad & \rho_{1A} = 3\,;\, \rho_{1B} = 1 \\ A \succ_{i_2} B\,;\, d_{i_2} = (0,1) \quad & \rho_{2A} = 2\,;\, \rho_{2B} = 3 \\ A \succ_{i_3} B\,;\, d_{i_1} = (0,1) \quad & \rho_{3A} = 1\,;\, \rho_{3B} = 2 \end{aligned}$$

Recall that the higher the ρ, the higher the priority. Normally we would add a lottery number but there is no need for it with this toy example.

We use the CAM algorithm to assign students their optimal fair matching. At the beginning, both schools start with a cut-off at 3, and at the first step, all three families apply to A. Because i_1 has top priority at A and is not feasible, we adjust the cut-off successively until the one at A reaches 0 and the daycare remains empty. At the same time at B, the different steps of the CAM algorithm leads us to a final assignment of i_2 while i_1 and i_3 are unassigned. The matching is thus $\mu = \begin{pmatrix} A & B & \emptyset \\ & i_2 & i_1,\, i_3 \end{pmatrix}$

Because i_1 is not initially feasible at her favourite daycare, she imposes a negative externality on the others and still ends up unassigned. Now, let us remove the daycares in the preference lists that are initially infeasible. The new preferences are given by:

$$\begin{aligned} B \succ_{i_1} \emptyset\,;\, d_{i_1} &= (1,0) \\ A \succ_{i_2} B\,;\, d_{i_2} &= (0,1) \\ A \succ_{i_3} B\,;\, d_{i_1} &= (0,1) \end{aligned}$$

Now let's run the CAM algorithm again with these preferences. The final assignment is obtained in 2 rounds with final matching $\mu' = \begin{pmatrix} A & B & \\ i_2 & i_1,\, i_3 & \end{pmatrix}$

In this example, removing initially infeasible demands from the preferences improves the assignment of all applicants. Now, i_2 gets her first choice and i_1 and i_3 are assigned to their second choice instead of being unassigned. This improvement comes at the cost of justified envy for applicant i_1, but this envy is based on an initially infeasible demand.

Proof of Theorem 1

Proof. Fix a problem $P = (I, S, B, \mathbf{q}, \underline{q}, \theta)$ and consider an applicant i. A policymaker defines rules for the DAM and chose a mechanism $\phi := P \Rightarrow \mu$ among CAM and KDA.

μ is the assignment using ϕ on problem P and by optimality of these algorithms, there are no stable assignment that are preferred by parents and respects the corresponding definition of envy-freeness.

Let $\succ_i$: $\{s_1, \ldots, s, \ldots, s_n\}$ be the reported preferences of i_i and let $\mu_i = s$. s can be any school in the preference list $\succ_i$ and also be $\emptyset$ if applicant i is unassigned.

Suppose that there is at least one initially infeasible demand in $\succ_i$ and let's call it s_k[80]. Let us denote $\bar{\succ}_i \equiv \succ_i \setminus \{s_k\}$ the new set of preferences without the initially infeasible school s_k.

We want to show that assignment μ with preferences $\succ_i$ is also (weakly) envy-free with $\bar{\succ}_i$. We prove by contradiction that the SOFA with initially feasible can only be better than the SOFA with full preferences. For this proof, we use fairness defined as envy-freeness and the CAM algorithm but the demonstration is the same with weak envy-freeness.

[80] Note that $s_k \neq s$. Indeed, if s_k is not feasible for i, then by feasibility of a matching with ϕ, i cannot be assigned to s_k.

Envy of applicant i in the matching with modified preferences $\succ_i^-$:

1. Suppose that when i is assigned to $\mu_i = s$ using the modified preferences $\succ_i^-$, she has justified envy.
2. Then $\exists\, s'\ s.\,t.\ s' \succ_i^- \mu_i = s$ and $\exists\, i' \in \mu_s'$ with $\pi_{i's'} < \pi_{is'}$.
3. Since $\succ_i^- \equiv \succ_i \setminus \{s_k\}$ where s_k is not feasible for i, then $s' \succ_i^- \mu_i = s \Rightarrow s' \succ_i \mu_i = s$.
4. Thus, if i has justified envy when assigned to $\mu_i = s$ using the modified preferences $\succ_i^-$, she also has justified envy when assigned to $\mu_i = s$ using preferences $\succ_i$. A contradiction with μ being the SOFA.

Therefore, the matching $\mu_i = s$ using the modified preferences $\succ_i^-$ is also envy-free.

Envy of other applicants i' in the matching with modified preferences $\succ_i^-$

1. Since only $\succ_i$ has been modified, no other preferences have been changed.
2. Therefore, if any applicant i' has justified envy with assignment μ with preferences $\succ_i^-$, she also has justified envy with assignment μ with preferences $\succ_i$. A contradiction with μ being the SOFA.

Therefore, the matching μ using the modified preferences $\succ_i^-$ is envy-free, completing the first part of the Proof.

Next we want to prove that this assignment is preferred to the assignment with the full set of preferences.

1. if μ is envy-free for the assignment problem P with $\succ_i^-$ instead of $\succ_i$,
2. then, by optimality of the matching of the SOFM, the outcome μ' of the CAM algorithm on the same problem P but with $\succ_i^-$ instead of $\succ_i$ is the student-optimal fair matching.

Thus, it is weakly preferred to μ and therefore Pareto dominates μ

QED

F.III Proof that our assignment mechanisms can be used for evaluation

Our general challenge is well summarised by Duflo (2017): "*Well-designed plumbing experiments can sometimes introduce variation that does not exist in natural conditions, and thus generate a counterfactual to illuminate theoretical mechanisms that are not easily observable in nature*". The econometrics of school choice is now mature enough to clearly define conditions under which we can use assignments using matching algorithm to measure causal effects on compliers[81].

One condition proved by Abdulkadiroglu et al. (2017) is that mechanisms need to satisfy *equal treatment of equals*. This condition is true if

1. the mechanism Φ is anonymous,
2. the distribution of probability of the lottery is symmetric

In other words, the mechanisms must not consider specific applicants differently but through priority rules and diversity quotas.

For every mechanisms we considered, we defined a lottery number ϵ_i that is $i.i.d.$ drawn from $U_{[0;1]}$. This random number is used to define the applicants' score π_{is} at schools s by combining with her priority level ρ_{is}. For each realisation of the tie-breaker, an algorithm returns a matching. For a given problem, a mechanism generates a distribution of probabilities over possible matchings, which is referred to as a **stochastic assignment**. A stochastic assignment generates a matrix $\mathcal{P}$ of size $|I| \times S$ where the entry p_{is} represents the probability that applicant i is assigned to school s s.t.

[81] For a recent review, see the chapter by J. Angrist, Hull, and Walters (2023) in the Handbook of economics of Education, 2023 edition.

- $\forall i \in I$, $\forall s$: $0 \leq p_{is} \leq 1$.
- $\forall i \in I$, $\sum_{s=0}^{S} p_{is} = 1$.
- $\forall s$, $\sum_{i \in I} p_{is} \leq q_s$

We have three main algorithms (and their associated soft quota version):

1. Assignment of families with SPDA soft quotas
2. Assignment of families using CAM
3. Assignment of families using KDA

for which only the definition of family types θ evolved. For problem 1, SPDA with soft quota satisfy ETE. Indeed, DA with lotteries has been shown to satisfy ETE by Abdulkadiroglu et al. (2017) and we simply break schools into seats and use SPDA. Hence the results of Abdulkadiroglu et al. (2017) apply immediately: we can compare families with the same propensity score.

For the other algorithms, we need to formally prove that they also satisfy ETE for the evaluation problem.

Assignment with days adds another component to family type θ which becomes more complex but still contains all relevant information for daycare assignment. As long as we use lotteries drawn from symmetric distributions of probabilities (which is the case for uniform distributions), the proof of Abdulkadiroglu et al. (2017) naturally extends to assignment with KDA and CAM, as shown in theorem 3.

Theorem 3. *The CAM and KDA algorithms satisfy equal treatment of equals.*

Proof. Fix a problem $P = (I, S, \mathbf{q}, \underline{q}, \theta)$, consider two applicants i and j of types θ_i and θ_j and denote ϵ_i and ϵ_j their lottery. Consider the mechanism $\Phi := P \Rightarrow \mu$ among CAM and KDA, with or without soft quotas. Denote $\mu_i(\epsilon_i) = s$ and $\mu_j(\epsilon_j) = t$ their assignments with the realised lotteries, which can be $\emptyset$ if they remain unassigned.

Suppose that individuals are of the same type $\theta_i = \theta_j = \theta$. Then $\succ_i = \succ_j$ and $D_i = D_j$ and $\rho_i = \rho_j$ and $g_i = g_j$ with soft quotas.

Since $\succ_i = \succ_j$, family i and j are ranked by the same schools (buckets) and their positions in these rankings depend on $\pi_i = f(\rho_i, g_i, \epsilon_i)$ and $\pi_j = f(\rho_j, g_j, \epsilon_j)$ where $f(\cdot)$ is common and defined within the algorithm. Since $\theta_i = \theta_j$, their rank only depends on the realised lotteries ϵ_i and ϵ_j. It is then obvious to see that if we swap their lottery and consider the ranking with $\pi'_i = f(\rho_i, g_i, \epsilon_j)$ and $\pi'_j = f(\rho_j, g_j, \epsilon_i)$, then we just swap family i and j ranks. Now, consider in particular the ranking at daycare s where i is assigned with $\mu_i(\epsilon_i)$. It is clear then that if we swap the lotteries, we swap the rank at school s_i and by design $\mu_j(\epsilon_i) = s$. Therefore Φ respects equal treatment of equals.

QED

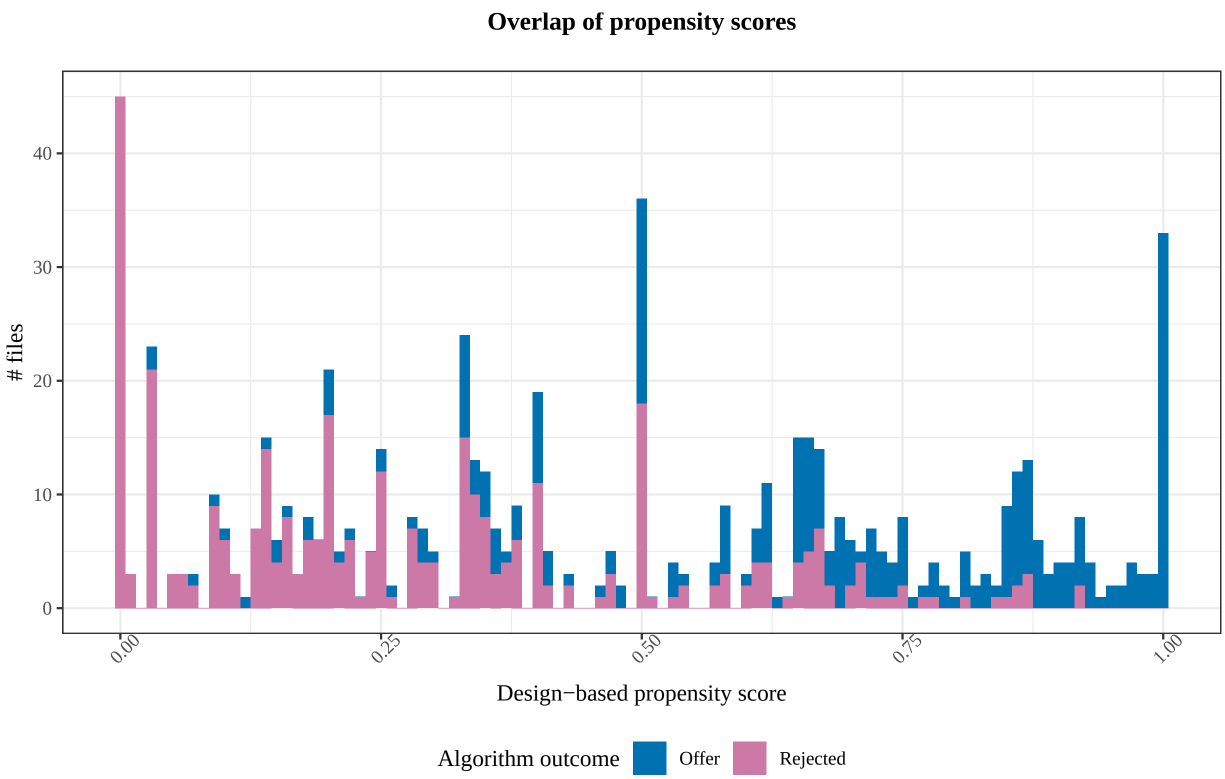
Our models provide the right setting for our evaluation problem.

Tie-breaking introduces artificial stability constraints since, after tie-breaking, schools appear to have strict preferences between students for whom they are indifferent. Abdulkadiroğlu, Pathak, and Roth (2009) showed that single tie-breaking, *i.e.* one random lottery per student and not per school, has superior welfare properties. Furthermore, there is no strategy-proof mechanism (stable or not) that Pareto-improves assignment with deferred acceptance with single tie-breaking. In practice, our random tie-break substitutes alternative criteria. Typically, assignment committees use application dates, the *quotient familial* or ad-hoc decision rules to select participants with the same priority level.

F.IV Randomness in the assignment: Illustration

In every round of assignment, we simulated a million alternative match and compute individual propensity scores as the frequency of assignment over these simulations. Over the four first rounds we implemented, 56.7% of all demands had a 0 probability of being assigned and 29.2% had a probability of 1. The remaining 14% were subject to randomisation and we present their distribution by offer in figure F.30.

Figure F.30: Distribution of propensity scores among those subject to random assignments



Sources: ISAJE, Case Study I − 2020 : 2023 − first rounds only.
Notes: We present the histogram of the propensity scores over bins of .01 for those with propensity score in]0;1[.

The distribution of design-based propensity scores for the instrument is neither smooth nor overlapping. This is important for the next steps of our work since a lot of econometric methods rely on well-behaved probability functions. Results can be sensitive to bin size and numbers, a regression function will typically put a lot of weight on bins with more observations, not where the propensity score is higher (Słoczyński 2022) and so on.

Part 2

Social investment in the labour market: the case of single mothers

Chapter 2.

Welfare-to-what ? Experimental evaluation of an activation programme for single mothers in poverty in France

I want to thank Marc Gurgand for his supervision and guidance over the years and Karen Macours for her advices and support. I am grateful to Camille Terrier and Anne Boring whose comments in the pre-defense really helped improve this paper. This project and associated researchs would not have been possible without the initiatives and maintained efforts of Gabriel André, Bernard Tapie, Florence Thibault, Virgine Gimbert, Lucie Gonzales. I owe a great deal to Alex Galitzine and Nirsynne Nahhal whose stimulating discussions, corrections and proofreadings greatly improved this paper. Special thanks to Pedro Sant'Anna for sharing codes to adapt the did package to this setting. I am grateful to Bruno Palier, Antoine Bozio, Robin Huguenot-Noël, Michael Zemmour, Clement Carbonnier, Saad Loutfi and the participants of the Labour and public policies seminar of Paris School of Economics, the Séminaire Travail en Économie Politique (STEP) of Paris I Panthéon Sorbonne and the informal research group on poverty at LIEPP (Sciences Po) for their stimulating comments questions and suggestions at various steps of this project.

Abstract

A quarter of households are single parents with their children, with mothers being the primary caregivers in the majority of these cases. This demographic represents a large share of the most vulnerable populations and is therefore the focus of anti-poverty policies, which now rely heavily on active active labour market policies. This study examines the causal effects of an intensive welfare-to-work programme targeting single parents on long-term welfare in France. In the experimental sample, 94% were women and 98% lived below the poverty line. The programme started in 2018 for three years and was extended for two more years because of the Covid-19 pandemic. It provided year-long intensive support, including individual and group sessions with highly trained social workers and childcare. The cost per participant was approximately € 2800, 4 times the usual spending for social support. Using a block-randomised encouragement design, 844 encouraged single-parents of 4 cohorts were invited to meetings presenting the programme, with 38% signing-up on average, whereas the 828 in the control group received no intervention. The analysis shows that participation increased from 28% to 48% from the 1st to 4th cohort, improved by more threatening invitation letters on the one hand, and more human enrollment through individual meetings in the programme's premises and testimonies from former participants on the other. Compliers are more likely to have less than a high school diploma, be among the poorest half, in their thirties, and registered with the Employment agency. The programme slows job finding rates during its first half, leading to a lock-in effect on poverty rates, disposable incomes, and employment. These effects swiftly fade-out and there are no average effect by the programme's end, for the following year, and up to 45 months since random assignment for the three first cohorts. Participants have higher employment rates than other comparison groups, but this difference is entirely driven by selection biais. The latter is so strong that estimates using the next-best identification strategy - modern doubly-robust difference-in-differences - fails to include experimental estimates in the confidence intervals. I provide additional analyses on distributional effect on incomes and treatment effect heterogeneity. These results are in line with the literature showing that welfare-to-work programmes and reforms' effects on impoverished single parents' employment are often weak, and strongly depend on a good buisness cycle. Yet to my knowledge, this is the only experimental evidence of the (absence) of benefits from welfare-to-work policies for poor single mothers in France. These results are therefore highly relevant to the political debates in France marked by a particularly violent conservative turn towards the most vulnerable.

- **JEL Classification Numbers** : I38, J16, J18
- **Keywords**: Welfare-to-Work, single parents, active labour market policy, long-term unemployment

Résumé

Les familles monoparentales représentent environ un quart des ménages avec enfants et sont principalement dirigées par des femmes. Elles font souvent face à d'importantes difficultés les plaçant au coeur des politiques de lutte contre la pauvreté, désormais largement centrées sur des politiques d'activation du marché du travail. Cette étude examine les effets causaux d'un programme d'accompagnement intensif visant à faciliter le retour à l'emploi de familles monoparentales, bénéficiaires longue durée des minima sociaux en France. Dans l'échantillon expérimental, 94 % sont des femmes et 98 % vivent en dessous du seuil de pauvreté. Le programme a débuté en 2018 pour trois ans et a été prolongé de deux années suite à la crise Covid-19. Il proposait un accompagnement intensif d'un an, comprenant des séances individuelles et de groupe avec des travailleurs sociaux hautement qualifiés et des possibilités de garde d'enfant. Le coût par participante était d'environ 2800 €, soit quatre fois les dépenses habituelles pour l'acompagnement social. En utilisant un recrutement basé sur un encouragement randomisé stratifié, 844 parents des 4 premières cohortes ont été invités à des présentations du programme, avec une adhésion moyenne de 38 %, tandis que les 828 des groupes témoins n'ont reçu aucune intervention. Les analyses montrent que la participation s'est accrue de 28 % à 48 % entre la première et quatrième cohorte, favorisée par un courrier plus menaçant d'un côté, et un accueil plus humain via des rendez-vous individuels sur les lieux de l'accompagnement et des témoignages de parents des cohortes précédentes de l'autre. Les participantes sont plus susceptibles d'être trentenaires, parmi les plus pauvres, les moins diplômées et déjà inscrites à pôle emploi. Dans les premiers 9 mois, le programme ralentit les reprises d'emploi, entraînant un effet de *lock-in* sur le taux de pauvreté, les revenus disponibles et l'emploi. Ces effets disparaissent d'ici la fin du programme, sans amélioration l'année suivante, et jusqu'à 45 mois depuis le tirage au sort pour les trois premières cohortes. Le taux d'emploi des participantes est pourtant plus élevé que celui de tous les groupes de comparaison. Mais cet écart relève uniquement du biais de sélection dû à des différences inobservables. Le biais de sélection est si fort que les estimations utilisant la meilleure stratégie alternative - des différences de différences appariées et doublement robustes - rejettent les estimations expérimentales des intervalles de confiance à 95 %. J'analyse également l'hétérogénéité de l'effet du programme et ceux sur la distribution des revenus potentiels. Au final ces résultats s'ajoutent à une littérature abondante et croissante pointant les effets faibles ou souvent délétères des politiques d'activations sur les familles monoparentales pauvres. Il s'agit toutefois à ma connaissance de la seule expérimentation aléatoire en France de programme d'accompagnement de familles monoparentales. Ses résultats sont donc particulièrement pertinents dans les débats politiques récents, marqué par un tournant conservateur particulièrement violent à l'égard des plus vulnérables.

- **Codes Journal of Economic Literature** : I38, J16, J18
- **Mots clés**: Welfare-to-Work, single parents, active labour market policy, long-term unemployment

I Introduction

Over the past fifty years, the share of single-parent households have become increasingly prevalent in many OECD countries, now representing around one-fourth of all households in the European Union (Nieuwenhuis 2020). This trend is associated with increased risk factors for both parents and children, such as financial insecurity, isolation, lack of access to quality healthcare and childcare, or poor health and housing conditions (Broussard 2010; Duriancik and Goff 2019). Welfare States have implemented various policies but their level of generosity or conditionality remains at the heart of major tensions, involving antagonistic views on social justice, family dynamics, and the role of work in society, all through the lenses of gender, social origins, and racial origins (Moller 2002; McLanahan 2004b; Reese 2005; Carcasson 2006; Foster 2008; Robert A. Moffitt 2015; Reingold and Smith 2012; Mangin 2021; Herbst-Debby 2022).

Since the 1990s, OECD countries have shifted towards active labour market policies (ALMPs), with distinct waves of adoption characterised by neoliberal then inclusive growth narratives, broadly advocated by economists (Peterson 1997; J. P. Martin 2015; Crépon and van den Berg 2016b; Peden 2017). Inclusive growth aligns with what Périvier and Sénac (2017) call a form of "*neoliberal morality*" through the *social investment paradigm* (Jenson 2010 ; Giuliano Bonoli 2011 ; Morel, Palier, and Palme 2012b). The latter try to reconcile economic efficiency with social justice concerns and shares with activation the idea that labour is the first protection against poverty. Social investments aim to both increase and use human capital to foster labour market participation through - among other things - education, training and job search assistance. However, it differs from activation by emphasising the idea that social inclusion cannot go without protection for the most vulnerable: the State must also play a role as a social buffer (A. C. Hemerijck 2014). Social investment had been an essential political framework in the European union over the past 20 years, leading Knijn, Martin, and Millar (2007) to argue that ALMPs have become the main policy framework for single parents.

While ALMP often have positive effects on average (Card, Kluve, and Weber 2018, 2010; Vooren et al. 2019a), they have been criticised for their impact on vulnerable groups like single parents and their children (Ellwood 2000; Campbell et al. 2016b; Avram, Brewer, and Salvatori 2018). For instance, the Cochrane systematic review of Gibson et al. (2018) summarises the effects of welfare-to-work programmes on single parents and reports almost no effect on employment, income or on mental health. Furthermore, and even when employment and income were higher for the lone parents in welfare-to-work, most participants continued to be poor. Despite large public spendings in activation policies, they often fail to alleviate joblessness and poverty among single-parent households, extending the risks to their children growing and remaining in poverty (Vandenbroucke and Vleminckx 2011b ; Løken, Lommerud, and Holm Reiso 2018; Rodríguez 2023).

In this paper, I evaluate the effects of an intensive welfare-to-work programme targeting single parents on long-term welfare in France. This programme called "*Reliance*" has been rolled out each year from 2018 to 2022 in the urban area of Nancy, the 16th largest urban area located in the North-East of France[1]. It targets single parents under 50 years-old registered for the French minimum income scheme: the *Revenu de solidarité active* (RSA). The initial stages of the support process involved diagnosing the main problems (such as over-indebtedness, housing, healthcare, and children's education), ensuring take-up of all social transfers and registration to the Employment agency. Throughout the year the intervention was organised alternately between group and individual sessions, at convenient times with regard to schools and daycare timetables. Participants involvement required about 15h/per week and, importantly, they could attend with their child(ren), for whom shared childcare duties among participants were organised in a dedicated space. The experimental sample is composed of 1671 single parents, 95 % women, 97% households living in poverty.

The *Reliance* programme is deeply rooted in the social investment paradigm by its target, policy objectives and design. It draws significant resources from various institutions[2], *investing* approximately €2,800 per participant - four times the usual spending per RSA recipient - in order to foster labour market participation and, ultimately, reduce poverty of highly vulnerable families. In the literature, these ingredients are usually predictive of higher effects on labour market participations. For instance, Bloom, Hill, and Riccio (2003) use data from three large-scale, multi-site random assignment experiments in the US and show that emphasis on quick return to employment,

[1] It was initially designed for three cohorts but two more were added because of the Covid-19 pandemic. See section II for details.
[2] Including the National family allowance funds and its local branch, the Caisse des dépôts, and the Departmental council.

personalised support, limited use of basic education and low staff caseloads increase effect size on employment and incomes. However the systematic review of Gorey (2009) shows positive effects are only found when the programme involve access to affordable childcare and decline when unemployment rates rise and jobs become harder to find. Conversely, the review of the effects of ALMP for women in Europe by Bergemann and Van Den Berg (2008) reports mostly positive effects, stronger for women than men, especially for training programmes. However, the vast majority of these research (39) are quasi-experiments while the only 4 randomised trials give opposite results (two positives, two negatives).

Relying on this state-of-the-art welfare-to-work programme for single parents in poverty, the main challenge is to ensure an evaluation that truly inform public policies. The analysis is thus guided by the following causal questions:

Does the programme increase labour market participation ? Does it reduce poverty ?

To answer that, I designed a staggered randomised experiment to assign encouragement each year using a block-random assignment defined by the product set of the number of children, unemployment registration and number of years on welfare. I focus on the four first cohorts of approximately 417.75 households each, with a total encouragement group of 843 households, and 828 controls. The former were invited to participate via formal letters, SMS reminders and phone calls. Throughout the years, the recruiting process has been adapted to foster participation. Letters were made more threatening, meetings were organised in the programme's meeting room, changed from collective information sessions to face-to-face individual meetings with project managers and former participants providing testimonies. On average, the take-up rate was about 38%, increasing from 28% to 47% between the first and fourth cohort. Participation is higher among the poorest, the least educated and those already registered at the Employment agency. I use matched administrative data from the National family allowance fund from January, 2017 to June, 2023, creating a panel dataset of 102749 observations to measure the effect of the programme on employment and poverty. I support and discuss the quantitative analysis with results from a qualitative evaluation conducted by an independent team of consultants (FORS 2020).

The results of this experiment are highly relevant to the French political debate, the past and forthcoming reforms. During its implementation, this programme was deemed very promising. It is praised in several official reports including those of the evaluation of the "Anti-poverty Strategy" launched by president Macron in 2018. It received the visit of the Minister of Health and Social affairs and a Secretary of State and quoted in MPs reports. However, social investment have been dropped from the government's narrative which, in a few years, took an unprecedented conservative turn[3]. After several waves of reforms tightening unemployment insurance eligibility and sanction measures, the French government[4] plans to enforce workfare obligations for all RSA recipients and already adopted several reforms towards it. As of January 2024, the Employment agency relabelled *France travail* is to oversee all RSA recipients by 2025, mandating 15 hours of work or social support under risks of monetary sanctions.

Contrary to this experiment, this new policy is more coercive and far cheaper. Yet this programme's effects can serve as an upper bound for the expected effects of the reform on poor single mothers. Indeed, There is a dearth of quality evidence on the effect of welfare-to-work programmes in France (P.-H. Bono et al. 2021). In a review of the effect of social support on various dimensions, Cervera et al. (2017) note just a handful quantitative analyses in France. The limited exceptions mainly focus on employment outcomes for the unemployed or young adults. A systematic review of the literature on the effect of ALMPs on long-term unemployed individuals between 2000 and 2015 by Abadia et al. (2017) reports only one study involving welfare-recipients of the former minimum income scheme (RMI), before the introduction of in-work benefits (Crepon et al. 2013).

This study presents a unique experimental evaluation and contributes to filling this hole in the literature on welfare-to-work policies for single parents in poverty. While it is always a researcher's duty to provide clear and high quality evidence with transparent empirical method, the stakes in this study feel notably higher. To uphold the research's integrity, all analyses and outcomes were preregistered, thereby mitigating potential political pressures and preventing cherry-picking. In addition, this article is knitted using Rmarkdown, and all codes generating results are embedded in the files to ensure replicability and transparency.

By design, random assignment identifies the intention-to-treat effects of the programme and the effects on participants when used as an instrument for enrollment and excluding other causal path between encouragement and

[3] For discussions on these shifts around the 2022 election, see for instance Knapp (2022),Hewlett and Kuhn (2022) or Durovic (2023)

[4] https://travail-emploi.gouv.fr/

outcomes. The staggered entry of cohorts and the treatment effect dynamics by time-to-event require careful estimation strategy. Staggered designs have been the focus of many recent methodological contributions[5]. I use stacked regressions with inverse propensity score weighting to estimate the intention-to-treat parameters, interacting treatment effects with relative months dummies and a full set of block $\times$ months fixed effects. These interactions saturate the model and ensure comparisons with clean controls (Sun and Abraham 2020; Freedman et al. 2023). As a robustness check and mean to compare experimental estimates with second best identification strategies, I adapt results of Callaway and Sant'Anna (2020a) to aggregate month-cohort specific treatment effects to this experimental setting. As for instrumental variables, I use results of Borusyak, Hull, and Jaravel (2022) on shift share IV and simply demean the instrument with the block-specific share of assignment and retrieve the treatment effect on the treated due to one-sided non-compliance. I use cluster-robust standard errors at the block level and control the family-wise-error rate using either the Holm–Bonferroni correction or wild-cluster bootstrap (Hothorn, Bretz, and Westfall 2008; Alberto Abadie et al. 2022; C. de Chaisemartin and Ramirez-Cuellar 2022; MacKinnon, Nielsen, and Webb 2023).

In brief, I find a strong lock-in effect for the first half of the programme that reduces employment on the treated by -10 percentage points, income per capita by about € 85 each month and slows the climb out of poverty. These effects dissipates by the end of the training and I find no average effect on employment and incomes for at least a year after the end of training for the four first cohorts. Over that period, only 10% of the sample earned more than the poverty line and the distribution of potential income per capita of treated and untreated compliers are very similar.

Patterns of treatment effect are heterogeneous across cohorts until the end of the programme, but this heterogeneity stems mostly from differences in the counterfactual. Covid-19 and the increase of in-work benefits may explain these short-lived differences. However, at the end of the programme, the effects on employment and incomes are homogeneous across cohorts. In addition, the treatment effects on incomes and employment for those who had one child and those who had three or more children are reversed and puzzling: For those with one child, the programme reduces employment but has no effect on incomes ; For those who had three or more children, the programme seems to increase employment but reduce incomes, although both are imprecisely estimated. These results suggests heterogeneous changes in the composition of disposable incomes between families of one and three children. The programme gradually increases total cash transfers among participants. However, this effect is mostly mediated by changes in household structures.

As-treated analyses show large selection effects and the next best identification strategy using matched difference-in-differences with the control group rejects the experimental estimates from the 95% confidence interval. I conclude that the programme attracted those with the highest employment potential but slowed their re-entry into employment (lock-in effects), inducing significantly lower disposable income until the end of the programme, and no effect after. Although the programme selected those with highest employment potential, labour market participation is always lower than 40 % and very few exit poverty.

These results are consistent with a large strand of literature showing that active labour market policies, especially workfare and welfare-to-work programmes may worsen the situation of vulnerable single parents (Smedslund 2006; Gorey 2009; Mogstad and Pronzato 2012; Brady and Cook 2015; Campbell et al. 2016b; Gibson et al. 2018; Avram, Brewer, and Salvatori 2018; Johnsen and Reiso 2020). Economics as a profession has a lot of responsibilities in the way society thinks of its problems. As Hirschman and Berman (2014) notes, "*the spread of economic discourse reshapes how non-economist policymakers understand a given issue. The spread of economists' technical tools determines the information available to policymakers and changes the process of decision-making*". Economists' narrative on welfare-to-work have been highly consensual despite many negative results. With more and more robust evidence coming-up with opposite effects from what policy makers intended, or economists thought would happen, perhaps the time is right for a paradigm shift.

The remainder of this research is organised as follows. Section II introduces the institutional setting and the intervention. Section III describes the experimental design, data and descriptive analysis. Section IV discusses the identification and estimation strategies. Section V presents the effects of the programme on the main outcomes and Section VI explores plausible mechanisms through a heterogeneity analysis. The interpretation, limits and policy implications of these results are discussed in section VII.

[5] See for instance the recent reviews by J. Roth et al. (2023).

II Supporting single mothers in poverty in France: Context and intervention

The programme targets single parents on long-term welfare in France through intensive social support. To understand the main difference between the treatment and the counterfactual, this section presents the general framework of active labour market policies for single mothers in France and discuss how this population fare under this regime. Then, I present the programme in detail and contrast it with the counterfactual and a workfare reform expected in 2025.

1 Single parents and the French minimum income scheme (RSA)

Before 2008, ALMP primarily aimed at reducing labour costs, particularly for low wages, and increasing employment rates through the sharing of working hours. The 35-hour work-week and the introduction of part-time and short-term contracts, as per Kramarz, Nimier-David, and Delemotte (2022), notably improved women's labour market participation, especially for married women. The pivotal turn in 2008 introduced a new welfare regime with in-work benefits and mandatory job search or social support, known as the Active Solidarity Income (*Revenu de solidarité active*).

The French *active* minimum income scheme (RSA): RSA merges the former minimum income scheme and the single-parent allowance, replacing the latter with the *RSA Majoré* for isolated parents for up to 1.5 years or until the youngest child is older than 3. Importantly, it includes two activation policies:

- "*RSA activité*," relabeled "*prime d'activité*" in 2016 (PA), is an in-work benefit inspired by the US Earned Income Tax Credit. It provides financial incentives for workforce re-entry through a monetary transfer, depending on household earned incomes, size, composition, and other social transfers.
- "Mandatory" social support and/or monitored active job search, managed by Departmental councils.

In 2020, around 2 million households received the RSA (Revenu de Solidarité Active) with a total cost of 12.6 billion euros (DREES 2022), roughly the same budget allocated to research, higher education, and innovation (13.4 billion euros in that year[6]).

RSA and PA are differential transfers with baseline levels assigned to eligible households, adjusted based on household structures and changes in earned incomes every quarter. Inspired by the Flat Tax model and the US earned income tax credit, the RSA design aimed to address the low incentives to work of the previous welfare system (Gurgand and Margolis 2008).

Labour market institutions underwent significant reforms in the Hollande and Macron presidencies, notably reducing both employment and unemployment protections (Milner 2017; Leruth 2017; Gazier 2019). In 2019, the French government increased monetary incentives following the Yellow Vest movement, widening income eligibility thresholds, increasing baseline amounts, and offering higher individual bonuses in the PA reform. This reform, estimated by Dardier, Doan, and Lhermet (2022), led to a 37% take-up increase and an average gain of € 70 per month.

Despite France's reputation for a generous Welfare State, 66% of RSA household live under the poverty threshold $i.e.$ less than € 1 140 per consumption unit[7], more than four times the general population's average of 14.8%. This population is highly isolated, with 49% not seeing friends or family in the previous month and 26% having no contact at all. Health problems are prevalent, with 21% reporting poor or very poor health, 43% having at least one chronic illness, and 38% stating it limits their capacity. Additionally, 22% are at risk of depression, and 15% had to forego seeing a doctor in the previous year due to economic reasons (DREES 2022, chap. 15 and 16). These statistics reflect the most recent reality, with documented deterioration since the COVID-19 crisis according to CNLE (*Conseil National de Lutte contre l'Exclusion*) (Duvoux and Lelièvre 2021).

[6] See the Senate report

[7] In 2019, the monetary poverty threshold computed as 60% of the median standard of living. See https://www.insee.fr/fr/statistiques/7710966.

Single mothers in France Single-parent families constitute 90% of beneficiaries of the "RSA majoré" (increased RSA), with a staggering 97% being women, and they make up 43% of RSA beneficiaries, 89% of whom are women. Female RSA recipients, particularly single women with children, face a monetary poverty rate of 73%, significantly higher than their male counterparts. The prolonged duration in RSA is associated with a decline in their standard of living, a trend worsening after the third year (Cour des comptes 2022, chap. 3).

Health issues, both physical and mental, often become a significant barrier to sustained employment or re-entering the workforce. Research by Delattre, Moussa, and Sabatier (2019), using French panel data, demonstrates a strong reciprocal effect of health status and employment, revealing a cycle where health deterioration at time t-1 predicts employment at time t, and past employment status influences health. RSA is sometimes used as a buffer allowance before accessing disability allowance, and the intersection of poverty, welfare eligibility, and disability is evident, with a quarter of new disability allowance recipients having previously received RSA benefits. However, recognizing their disability is challenging due to a lengthy process and unclear eligibility criteria, trapping them in a situation that hinders compliance with welfare-to-work obligations, leaving them in an uncertain status between the two institutions (Cattoen et al. 2022).

Moreover, the life trajectories of single mothers are often marked by traumatic experiences, including exposure to violence, domestic violence, and sexual assaults, extending back to their childhood. Studies, such as the analysis of data from the VIRAGE survey by E. Brown (2020) in France, reveal a high prevalence of violence, with its effects persisting post-separation. Notably, women no longer in relationships report experiencing harm more frequently, with a significant proportion enduring severe violence, often rejecting conventional marital norms and embracing enduring single parenthood.

Despite many being employed or having worked in the past, victims of violence often face higher instances of unemployment or inactivity for periods exceeding six months. Sometimes, Intimate partners use violence and coercion (IPVC) to obstruct women's economic independence[8]. Such violence contributes to increased vulnerability, self-deprecation, and even depression, leading to pronounced social isolation. Consequently, these women often feel illegitimate in returning to employment, especially with limited work experience.

2 Intensive social support for single parents on long-term welfare

The **Reliance** programme is an experimental comprehensive support initiative to foster "*sustainable reintegration into employment and society*" of single parents on long-term welfare[9]. It was funded and implemented by the Departmental council, the *Caisse des allocations familiales* (CAF) of Meurthe-et-Moselle, the National family allowance fund (Cnaf), and the Caisse des Dépôts. The programme targets eligible population residing in the urban area known as *Grand Nancy*[10], which encompasses 20 municipalities[11] and includes 8 neighbourhoods categorised as deprived for a total population of 420,120 inhabitants in 2019, of which 47,799 lived in single-parent families (INSEE 2023).

[8] For instance, Riger and Krieglstein (2000) report that, within a job programme in the USA, 47% of women who were victims of violence mentioned their intimate partners attempting to prevent them from pursuing education or training. Both victims and non-victims in this sample were discouraged from working by their partners. Spencer et al. (2020) analyse the effect of TANF for victims of IPV and also find such effects. For a review of the links between violence and family type, see Tur-Prats (2019).

[9] A web-page presenting the programme is still active in February 2024 : https://www.arelia-asso.fr/index.php/25-canevas/newsletterlirelasuite/510-reliance-est-un-nouveau-dispositif-experimental

[10] It is located in the heart of the Grand-Est region, which had 1,918,000 salaried jobs by the end of 2020. The employment dynamics of the region are mixed, with a net loss of 50,000 jobs from 2010 to 2020, primarily driven by declines in the industrial (-13.9%) and construction sectors (-8.1%). The tertiary sector remained relatively stable.

[11] Art-sur-Meurthe, Dommartemont, Essey-lès-Nancy, Fléville-devant-Nancy, Heillecourt, Houdemont, Jarville-la-Malgrange, Laneuveville-devant-Nancy, Laxou, Ludres, Malzéville, Maxéville, Nancy, Pulnoy, Saint-Max, Saulxures-lès-Nancy, Seichamps, Tomblaine, Vandoeuvre-lès-Nancy, Villers-lès-Nancy.

A welfare-to-work programme in the social investment paradigm The programme takes places in a renovated building in Vandœuvre-Les-Nancy, on the main tram-line at the heart of a neighbourhood prioritised for urban policy, equipped with offices for interviews, meeting rooms for group sessions, a children area, a communal kitchen, and computers. It is a delegation of a public service mission to private but non-profit operators: three well-established local associations. This experiment thus shares some common features with researches that compare public and private provision of workfares (e.g. Behaghel, Crépon, and Gurgand 2014). Social support is implemented by three experienced[12] social workers and an executive, for 82 participants per year, on average. This caseload per social worker is thus very low; for comparison, Jacquey-Vazquez (2017) reported that in 2015, The Employment agency intensive support advisors were each responsible for an average of 108 unemployed individuals. Participants enroll in the programme for a year hoping to find adapted and sustainable solutions to employment and precariousness issues, making the concept of social investment central to this experiment. The assumptions regarding the effects of the programme are based on both a *capacity-building* approach – where participants benefit from the programme by developing or maintaining skills, building a network, etc. – and an *emancipatory* approach that seeks to alleviate the specific burdens and obstacles faced by each family.

The support is based on a comprehensive assessment of beneficiaries' situations (personal, family), professional backgrounds, training needs, and environment. The initial stages of support involve diagnosing the main issues the family faces (overindebtedness, housing, health, children's education, etc.) and ensuring access to rights by resorting to national and local aids, including visits of social workers from the Family allowance fund. This step also aims at overcoming peripheral barriers to employment such as facilitating access to childcare for the families involved and confidence in their level of social transfers if they take a job. One objective is to quickly alleviate individuals' mental load, enabling them to regain their own resources (Anandi Mani et al. 2013; Schilbach, Schofield, and Mullainathan 2016a).

In this project, evaluation is a central component and has been conceived as part of the programme itself. Beyond my research, the Departmental council funded a qualitative analysis (FORS 2020) and coordinated two masters dissertations: one in sociology on how social workers perceive innovations in their work (Mahdi 2021), and one on non-take-up of social programmes (Chachou 2019). We also conducted surveys in order to measure the effects of the programme on more subjective dimensions[13].

The social investment perspective legitimises public intervention through a demonstration of its social returns in the medium to long term (Périvier and Sénac 2017). The programme would be considered a success if it generated positive "social returns" through "avoided costs" in social benefits. In this regard, the average cost per participant is estimated to be around € 2800, approximately four times the average expenditure for regular support (Mahdi 2021).

A highly promoted innovation Recent official reports have been highly critical of the implementation of job search monitoring and mandatory social support for RSA recipients (Pitollat and Klein 2018, Aout; Damon 2018; Cour des comptes 2022). Concerns include inadequate support reaching eligible recipients, a time-consuming referral and support process, heterogeneous content, and insufficient employment outcomes. This experiment takes a totally different path with intensive support, innovative recruiting process and a strong emphasis on measuring its effects. Specifics dully noted in the aforementioned reports, who cite this experiment as a promising way forward.

It also received significant political and media support, including an official visit[14] by the Minister of Solidarity and Health, Agnès Buzyn, and the associated Secretary of State, Christelle Dubos. The programme was awarded the Afigèse[15] 2021 Prize in the "public policy evaluation" category, despite the fact that the results of the impact evaluation had not yet been released. It has been featured in several articles in regional newspapers, radio broadcasts, and other specialised media outlets.

[12] They all have a master degree and several years of work experience.

[13] The objective is to measure relatively long-term outcomes and taking into account funding constraints, we decided to survey the cohorts at the same date and to repeat the questioning respectively 2 and 3 times for the 2019 and 2020 cohorts to see if these results evolve. The analysis of the responses to these surveys will be the subject of another publication.

[14] Meurthe-et-moselle.fr Actu - visite ministre

[15] *Association finances, gestion et évaluation des collectivités territoriales*, a network that brings together professionals in local public finance and management, public policy evaluation, and territorial public management. Press release

A highly vulnerable target population The target population is highly vulnerable. At the time of random assignment, the average standard of living is about 718 $€_{2015}$ and only 2.8 % are above the poverty line in 2019. In the qualitative evaluation (see details in section III), most single parents report health problems, employability constraints, budgetary difficulties, and housing instability. Often less educated and engaged in precarious employment, particularly those with migration backgrounds, they often withdrew from the labour market after the first child's arrival, entering RSA. They now juggle domestic responsibilities and parenting alone.

Facing economic hardship and social isolation, compounded by the dual stigma of being RSA recipients and single parents, they often experience a sense of "social shame," exacerbated in cases of prior violence. The qualitative evaluation underscores their vulnerability to life's setbacks, lacking social and economic safety nets, and being more exposed to health and family-related risks. Some had their children removed from custody or are under monitoring from children protection services. Several experienced intimate partner violence and coercion, which, for some, led to depression and post-traumatic stress disorder of the victims and sometimes their children. In this sample, only 20% receive child support from the other parent of their children.

Recognising the unique challenges of this population, Reliance provides comprehensive support, combining individual and group assistance. Thematic workshops cover diverse topics, involving local community partners, field trips, and leisurely outings. Workshops address benefit rights[16], parenthood, and digital literacy, aiming to humanise administrative procedures and alleviate emotional burdens. They offer practical assistance with applications for social housing, affordable school lunch, and suitable childcare options, as well as strategies for daily life organisation. Additionally, workshops explore self-awareness, relationships, and gender norms.

Complementary workshops cater to individuals' interests and backgrounds, including activities related to specific fields, outings, and well-being. Support days alternate between group and individual sessions, accommodating participants with or without children. Children have a designated area and are supervised by participating parents. The schedules align with school hours and participant availability, evolving based on observations and needs. Activities primarily focus on creating and validating realistic professional projects, addressing steps like education, internships, and improving job search efficiency, aligning expectations with job opportunities.

A staggered roll-out marked by the Covid-19 crisis To take the social investment perspective seriously, estimates of the causal effects of the programme are required. At the onset, we wanted to recruit a 100 participants every year, and policymakers defined the *success* of the programme as a 10 percentage point increase in employment. I conducted simple power computations with an *optimistic* take-up[17] of .5, 12 % attrition, mean employment change in control group of 10p p (SD=.3), 80 % power and 5 % two-sided test, reaching approximately 1200 households. We started with a target of three cohorts of 400 households.

Each cohort was randomly sampled and assigned to treatment. In general, caseworkers refer welfare recipients to such programmes, making comparisons with untreated recipients unlikely to recover causal effects of interest. This recruitment process is thus very different from common practices. In fact, it changes the pool of participants and reaches welfare recipients who would not have come otherwise. The qualitative evaluation strongly emphasise this results and documents a swift swing of opinion regarding random assignment from previously sceptical social workers. Surveying participants and non participants, FORS (2020) also reports a strong support from participants themselves.

Through this period, the economic environment was affected by the COVID-19 pandemic, and the increased in-work benefits in the 2019 reform of PA. The pandemic disrupted the implementation of the programme for the 2020 cohort and the economy was almost entirely shut-down for a while, before slowly recovering[18]. As a consequences, we opted for the pursuit of the programme and secured funding for two additional cohorts. We enlarged the sample for the 2021 cohort to increase precision. In 2022, the pool of eligible families that had not been already sampled was too small, especially for long-term recipients. To build the 2022 cohort, we sampled the remaining eligible population and random samples from the control groups of the previous cohort. From November 2021, composition of the control groups of the 4 previous cohorts change. In the end, the intervention has been rolled-out from 2018 to 2022 in a staggered design summarised in Figure A.11 in the Appendix.

[16] This component of the programme turned-out to be very important as shown in Galitzine and Heim (2024).
[17] This parameter was provided by project managers who expected a large enthusiasm from participants.
[18] We give more details of the adaptation and consequences of the pandemic in Appendix A.II.

III Data, experimental design and descriptive statistics

1 Data sources and main variables of interest

Matched administrative records The design of this experiment is based on repeated draws of random samples of eligible households from administrative records of the Departmental council. These samples constitute cohorts for which treatment has been randomly assigned following the protocol described in subsection 2. The initial datasets have been supplemented with these design variables and matched with monthly administrative records from the National family allowance fund (CNAF). The ALLSTAT files describe the situation of every beneficiaries for a given month. They contain information on the "*household heads*" and their possible spouse (including gender, year of birth, marital status, activity status, nationality, and so on), their dependent children (years of birth, alternated residence status for family benefits, absence of a parent, the legal benefits they receive (the individual social action aids they have benefited from), and several measures of household incomes. Details on the data sources and variables used in this analysis are presented in the Appendix B. Codes to generate these variables are available upon request. The main drawback is that they are not meant to measure employment but record relevant informations to compute all social transfers. However, they provide very good quality measures of incomes for the poorest and are consolidated over 6 months to account for treatment delayed, controls and adjustments.

I construct a baseline database of 2662 households from five cohorts measured the month before random assignment, including 548 households excluded from the experiment and 53 from the pilot study (see section 2 below). The *experimental sample* consists of 2073 households. Additionally, panel data from the National family allowance fund and data quality assessments reveal complete information for 92% of the post-randomisation period, with only 4.1% files lost, showing no differential attrition between treatment arms.

Qualitative evaluation FORS, a consulting firm specialised in social sciences and public policy evaluation *to la française*[19], conducted the qualitative evaluation of the programme under the sponsorship of the Departemental council, with the report finalised by the end of 2020. Throughout my analysis, I incorporate insights[20] from this report to support and contextualise my findings, referencing it as the qualitative evaluation or citing it as FORS (2020). The methodology was organised as follow: Initially, the mission was framed through meetings with programme pilots and the Reliance team, alongside documentary analysis. Then, the framework and survey tools were validated to ensure alignment with research objectives. The investigation unfolded in two main phases: the first involved interviews with participants from Cohorts 2018 and 2019, as well as focus groups and team meetings. The second phase expanded the scope to include follow-up interviews with Cohort 2019 participants, interviews with Cohort 2020 participants, and discussions with individuals who declined programme participation, along with interviews with partner organisations. The final phase encompassed the synthesis of findings and recommendations, including additional interviews with Cohort 3 participants and a collective interview with the Reliance team. Adjustments were made due to the COVID-19 pandemic, necessitating alternative data collection methods such as video conferences and telephone interviews, while still adhering to the research objectives and participant preferences.

Main variables of interest Households are identified with a unique id[21] and the cohort to which they belong. The Department's database contains the blocks' identification and associated variables[22], the encouragement (Z) and treatment (D) status, the date of randomisation from which we construct time-to-event variables. I also keep the Department's social workers' assessments (favourable or reserved).

[19] See Delahais and Lacouette-Fougère (2019)
[20] Translations of direct quotes from interviewees are my own.
[21] Cnaf does not have a national id for each household, identification is specific to the CAF families that are registered. When they move to another county or their relationship status changes, the household ID changes. I retrieved households that moved using other identification variables in the main information system.
[22] Number of children (1, 2, 3+), registration to the Employment agency (True/False), and length of time receiving RSA (2 to 5, 5 to 10, 10 or more years) ;

The main outcomes have been defined in the registration plan[23]. Table B.4 in the Appendix describes how their were constructed and alternative measurements used as robustness checks. This paper focus on two main outcomes in line with the main research questions:

- **Employment**, measured by reported labour incomes
- **Poverty**, measured by disposable incomes per consumption units and compared with the poverty threshold.

In the pre-registration, I also registered welfare and in-work benefit eligibility and total social transfers. Their analyses are presented in the appendix together with alternative measurement of the main variables, and briefly discussed together with the main results. In Galitzine and Heim (2024), I also investigate the effects of the programme on family structure (number of children and relationship status) and source of incomes.

Measures of labour market participation The main outcome of interest is labour market participation although I can only observe self-reported monthly household incomes used to compute social transfers. PA and RSA depends on all labour incomes in the households and other sources of incomes (including other social transfers) in the previous quarter. Each household must report their incomes by member of the household and type of income. I use these quarterly income reports and define employment as a dummy that equals one when the parent declares *her* positive labour income in a given month and zero otherwise. The main limitation is that I can only observe employment for families who report their quarterly incomes. Employment levels may be underestimated if those who stopped reporting their income did so because they earned more than the eligibility thresholds. As discussed in the previous section, these variables suffer from minor attrition which does not differ between the encouraged and control groups. Therefore, employment may be measured with error, but should not affect treatment effect estimations.

The database also contains detailed information on parents' "main" occupation, which can be classified according to the Labour force survey definition in 7 categories[24] and is also observed for parents who did not report their quarterly incomes. It has less missing cases but is less reliable nonetheless. Indeed, this variable is filled by CAF agents and there may be important delays, depending on controls or updates for motives unrelated to employment. Participants met with CAF social workers several times at the beginning of the programme and their files were likely updated while the control group was not. Those who do not report their quarterly incomes are even less likely to be updated on the right time. I only report estimates based on this variable as a robustness test in the Appendix.

Estimating the effect of the programme on poverty The Family allowance fund computes income per consumption units, which encompasses every source of income (labour and capital income, unemployment insurance, health pensions, and cash transfers from CAF) from every member of the household weighted by the number of consumption units. The first adult is weighted 1, any other member of the household that is above 14 years-old is weighted .5, and children under that age are weighted .3. Single-parent households receive .2 additional points. Note that this scale tend to overestimate the living standard of single parent households for whom economies of scales do not existent *a priori*. A 14 year old child is weighted the same way a partner is (Le Pape and Helfter 2023). In this research, I compare this variable with the poverty threshold from 2019, which is € 1140 according to INSEE. I convert incomes and threshold to 2015 values and estimate the effect of the programme on the share of households with less than this threshold to measure the effect on poverty.

[23] Note that in the registration plan, there are additional outcomes based on surveys which have not been exploited yet. My conclusions are based on this first set of hypotheses and the others will be analysed in another research, also involving data from the 2022 cohort and longer observation windows for the early cohorts.

[24] Other Inactive, Other Unemployed, Retired, Student Or Training, Unemployed, Unknown, Works

Covariates used in the analysis Blocking variables already accounts for the cross effects of number of children, registration at the Employment agency and years receiving RSA across cohorts. A large set of attributes is observable, further increased by panel data and the possibility to use past outcomes and characteristics as covariates. However, I remain parsimonious and only use a selected set of observables that correspond to the typical variables social workers would observe. The definition of these variables is presented in table B.5 in the Appendix. Following S. Athey and Imbens (2017b), I only use *dummified* versions of the covariates to keep a clear interpretation of all coefficients and limit bias from parametric restrictions.

Baseline outcomes typically predict subsequent outcomes and I include them measured the month before random assignment and centred. I use typical socio-demographic variables including dummies for French citizenship, quartiles of age, having children under 2, having children between 3 and 5 and having children older than 16. French citizen and younger worker usually face less constraints on the labour market (Hargreaves 2015; Anne et al. 2019), while children under three require intense care, preschool is mandatory from age three which may ease parents re-entry in the labour market (Goux and Maurin 2010). Older children may help care for younger and be more autonomous in general, thus possibly reducing parental constraints. Data from the Departmental council also include a measure of education that is known for approximately 80.5% of the experimental sample. I construct a simple High/low/unknown education variable, considering high school diploma and above as high education. To control for incomes, I use a dummy for taxable incomes 2 years before higher than the median[25], quartiles of disposable income per consumption unit at baseline, and dummies for housing benefits, public alimony (ASF) and receiving child support[26]. Finally, I define dummies for being re-sampled in the 2022 cohort and being assigned to the encouragement group, and allow specific trends for these groups.

2 Selection of participants

The design uses simple random sampling from administrative data among the eligible population on the one hand, and block-random assignment of encouragement on the other. This section discusses the recruitment process.

A) Sampling, pre-selection and random assignment

Each year, the Departmental council draws from its administrative databases a random sample of about 500 households that meet the following inclusion criteria:

- Single-parents under 50 receiving RSA and with at least one child present in the housing ;
- Resident of one of the municipalities of the *Métropole du grand Nancy*;
- Last RSA registration older than two years;

We ran a pilot study to test the programme's attractiveness and randomly selected 53 households in November 2017 and invited 41 of them to participate in the programme. Among them, 15 enrolled, giving us an idea of the expected take-up ($\approx$ 36.6 %). These families started the programme in January 2018 for a few months in a very small group before being joined by the participants of the 2018 cohort in April 2018[27].

[25] Which means no taxable incomes.

[26] I do not use early childhood allowance which is colinear with children under 3 and Family allowance and Family supplement which are colinear with number of children. Indeed, Family benefits are open from the second child, and the Family supplement from 3 children.

[27] Figure B.14 in the appendix shows the share of single parents with positive labour income in the pilot and the 2018 average. This first very small group had spectacular outcomes showing up very early. However, they lost their jobs during the pandemic and did not catch up since.

Caseworkers' orientation: A preselection After extracting the 500 files, civil servants of the job-support bureau of the Departmental council assessed files' *appropriateness* for the programme. They contacted each household's caseworkers and/or consulted their files to sort into three categories.

- "Favourable" are parents who would typically be referred to by their caseworkers for the programme had we used a more conventional recruiting approach.
- "Reserved" are generally formulated for families who have little contact with their caseworkers or if the known problems of the family make participation in such a programme *deemed unlikely* (typically because of known social or health problems)
- "Unfavourable" are given to families who already have a job or are already enrolled in another programme, to families followed by the child welfare service or to those for whom the programme is deemed inappropriate taking into account the information known to the services

The latter group represents an average of 21 % of the sample and were excluded from the experiment. The experimental sample is composed of households with favourable and reserved assessment. This pre-selection has an important consequence: the experimental sample is no longer representative of the population it was drawn from. At the same time, it is closer to ecological conditions and enables comparisons between those who would have been referred to the programme and those with *reserved* initial assessment, and how the treatment effect varies between these groups.

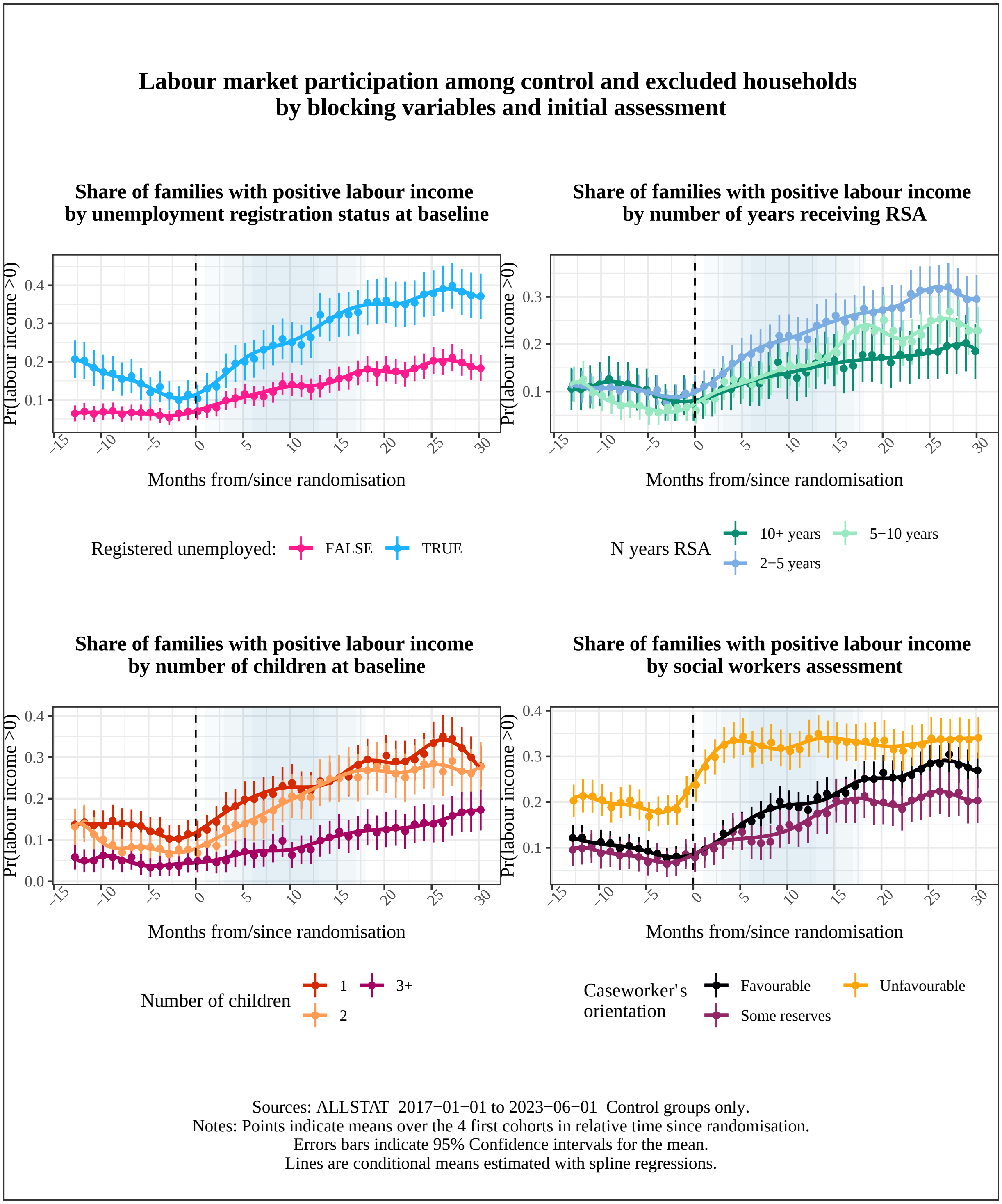
Randomisation protocol Each year, the Departmental council provided the random sample of eligible households with social workers' assessment and blocking variables. To assign encouragement, I use block-randomisation[28] within strata based on the cross product of:

- Registration status to the Employment agency (True/False)
- Number of children at home (1, 2, 3 or more)
- Years receiving RSA (2 to 5, 5 to 10, more than 10 years)

The product set of these variables define single parents' *type* between which I expect different reaction to the encouragement, different outcomes and possibly heterogeneous treatment effects. Registered unemployed are expected to be closer to the labour market, the number of children increases constraints due to parental obligation while it has been shown that the longer people receive RSA, the less likely they are to find a job (Cour des comptes 2022). These hypotheses are confirmed in the data, as shown in Figure III.1. Each panel illustrate the share of families in the control group with positive labour income based on the values of the blocking variables. It also shows the gap by social worker assessments including the excluded ones. The data are presented over relative months from randomisation, displaying mean values with 95% confidence intervals and smoothed conditional expectations estimated through splines. The key insight drawn from this visual analysis is that prior to random assignment, most groups exhibit comparable outcomes, and as time progress, the proportion of families with positive incomes increases but evolves distinctly across the different groups. Those registered at the Employment agency are significantly more likely to work before they are sampled and have a higher job finding rate and employment level than those unregistered. Files that were excluded from the experimental group are more likely to work prior to random assignment, and increase by more than 15pp in 6 months before reaching a plateau at about 1/3. Those with favourable assessments have a slightly higher job finding rate than those with reserved assessments.

[28] I draw and sort lotteries from a uniform random variable using Stata and treat the top first half, rounded above. Blocks' sizes vary within and between cohorts and the share of encouraged families is not exactly .5 in every stratum.

Figure III.1: Blocking variables and caseworkers initial assessment predict heterogeneous employment trajectories



Balance check and descriptive statistics Table B.6 provides the means, standard deviations, and differences between the encouraged and control groups. As expected, the random assignment of encouragement successfully balanced the groups across all variables used in this analysis. It provides the average values of the main outcomes and covariates at the time of random assignment. The sample is composed of 95% women aged 36 on average, 80% are French citizens, living with €$_{2015}$ 710 per month per consumption unit. The average household income is about

$€_{2015}$ 1400 while the average amount of social benefit is $€_{2015}$ 1300. About 1/3 has a child below 2 and 1/3 have a child older than 16. Half of them have a high school degree or higher while we don't know the education level of 1/3 of them. Only 1/5 receive child support from the other parent and 65% receive the family support allowance instead.

B) Increasing efforts to foster participation

At the beginning of the experiment, we were unsure whether we should foster participation through incentives and nudges or resort to more sever approaches, including threats of sanctions and benefit reductions. In the first year, the encouragement group received an official letter with information on the programme emphasising the potential benefits of participation and inviting them to public meetings where they could obtain more information. Following these meetings, interested families were invited to an individual appointment to sign a contract. Throughout the experiment, enrolment was voluntary, and parents could refuse to attend the meetings or enter the programme after the meeting. The take-up in the first cohort was the lowest[29]: 27.9 %.

We decided, in accordance with the international review board, to adapt the recruiting process for the subsequent cohorts. In particular, we opted for an ambiguous yet slightly more threatening letter as a way to foster attendance to the information meetings. From 2019 onwards, the invitation letters were sent earlier adding a bold sentence to the letter in French (see the template A.10 in Appendix A.I):

"I hereby inform you that **your participation in this meeting aligns with your obligations under the legislation of the RSA programme.**"

Our goal was to attract as many participants as possible to the informational meetings (by imbuing them with a perceived sense of obligation). However, the meeting place was moved from the Departmental council to the programme's meeting room, allowing prospective participants to visit the place and receive a warmer welcome[30]. The meetings retained their role of disseminating information and emphasising the benefits of the programme.

Moreover, former participants attended these sessions to share their experiences. Indeed, the content of the meetings also aimed to mitigate mistrust, diminish stigmas, and dispel any ambiguity regarding the mandatory nature of participation. Attendance at the meetings increased as a result, leading to a significant rise in the participation rate, reaching 36.6 %. At that time, the preliminary results of the qualitative evaluation revealed that the main motivations for participants was actually the fear of sanctions, and not potential benefits. In focus groups of participants, all agree that "*people would not come otherwise*", but stress the need for the programme to remain optional. An information only made clear during the meetings. Almost all individuals interviewed by the qualitative research team perceived their attendance as obligatory, or at least recognised a risk of their benefits being cut[31]. While the number of people who actually attended the collective meetings was sufficient, only 66 %[32] followed through with the programme later on. The project leaders estimated that an invitation to an individual interview organised by one of the social workers

[29] Note that compared with other comparable experiments in France such as Crepon et al. (2013) and Castell et al. (2022), this participation rate is rather high.

[30] A qualitative study was conducted in 2019 and 2020 surveying single parents of the three first cohorts (FORS 2020). It reports that these changes seemed to promote participation and strengthened the beneficiaries' commitment, both at the ontset and throughout the programme: the friendly environment, cleanliness and suitability of the premises, their aesthetics, etc. Participants emphasise material details that create a human and warm environment (such as coffee, for example), as well as more intangible aspects such as listening, kindness, and the "non-condescending" and non-judgmental attitude of social workers during these meetings. These elements form a framework that, according to participants, sets it apart from other social or support schemes they encountered in the past.

[31] For instance:

- "*I had received a letter. It was an information meeting, it was a requirement due to my RSA, I was obligated to attend, we had an attendance sheet, to say 'I came and I sign.' And afterward, it was our choice to stay or not. For me, I understood it that way.*" (Female, 44 years old, 2 children, on RSA since 2011)
- "*When I read the letter from CAF, I thought it was mandatory to do it, it was poorly presented. Otherwise, I thought I would lose all my RSA. It was poorly expressed, I think.*" (Female, 41 years old, 3 children, separated for 2 years, on RSA since 2009)
- "*We received a letter from the Departmental council, stating that it was a programme for single-parent families on RSA. It was not an obligation, but behind it, we knew there was a trap, your RSA could possibly be affected.*" (J., 39 years old, 3 children, on RSA for over 5 years).

[32] To comply with the General Data Protection Regulation, attendance sheets were only used by project managers to later contact participants who wanted to sign-in and destroyed after recruitment was complete.

of the project and former participants would be more likely to increase the participation rate by better explaining the interest in participating in this programme. Additionally, collective information sessions made some individuals uncomfortable, leading them to opt out. Individual interviews were expected to reduce judgement, comparison and stigma, a phenomenon often observed in the first waves, sometimes resulting in the feeling that the programme is not suitable for them. To address this issue, we changed the recruitment process from public sessions to face-to-face meetings with social workers from the programme and former participants. They provided testimonies and feedbacks of their own experience in the programme. This change in the recruiting process has been approved by the IRB and motivated by the fact that seeing the pool of other welfare recipients promoted stereotypical views and discouraged some from registering (FORS 2020).

Individuals who have experienced similar situations are perceived as more credible models than "institutional" voices in stimulating change among their peers: they have a better ability to establish empathetic relationships, convince others, and serve as role models. Mrs. D., a participant from the 2018 cohort and a "witness," explains: "*We realised that when it's presented by a former participant (me), it's perceived differently. It's amazing, this difference, the listening is different, we quickly noticed it.*" Furthermore, using the testimony of former beneficiaries also allows them to be valued and highlights their skills and experiences. For instance, a participant from the 2019 cohort explains: "*She explained her past to me, honestly, hats off. And you feel less alone. We're not alone in this case. And there's no judgement, we feel like we're all equal. I felt welcomed. It gave me a boost, that's what I needed in the end*" (F, 48 years old, 4 adult children, on RSA for 3 years, met in November 2019 during a mobilization day). Another says: "*It gave me a little more confidence to have someone who came like me, it reassures me*" (F, 34 years old, 4 children, met during the mobilisation day).

This change induced increased recruitment costs and duration, but ultimately led to a higher take-up rate: for the 2020 and 2021 cohorts, the take-up was 42.1 % and 47.2 % respectively. The 2022 cohorts received the same letter than the previous cohorts but did not receive text message reminders. The take-up rate was then 37.4 %. In the 2022 cohort, some participants did not received the reminders, probably explaining the lower take-up.

Table III.1: Average effects of encouragement on participation by cohort

	sample					
	Full sample	*Cohort 2018*	*Cohort 2019*	*Cohort 2020*	*Cohort 2021*	*Cohort 2022*
Encouragement	0.386***	0.279***	0.363***	0.421***	0.472***	0.378***
	(0.018)	(0.027)	(0.027)	(0.039)	(0.045)	(0.036)
Num.Obs.	2073	395	397	386	493	402
R2	0.282	0.195	0.249	0.300	0.354	0.266
R2 Adj.	0.249	0.156	0.214	0.266	0.329	0.228
Std.Errors	by: strataXc	by: strataXc	by: strataXc	by: strataXc	by: strataXc	by: strataXc
FE: strataXc	X	X	X	X	X	X

* p < 0.1, ** p < 0.05, *** p < 0.01

Sources: Sample: Cohorts 2018 to 2022 at the time of randomisation.
OLS regressions of participation on encouragement, recentered by within-block propensity scores and blocks x cohort FE. Cluster robust standard errors adjusted by blocks x cohorts in parenthesis.

Table III.1 summarises the average effects of encouragement on participation and separates estimations by cohort[33].

[33] Simply regressing participation on encouragement and block fixed effects

The take-up dramatically increased over the years and there are 19.25 pp difference between the first and fourth cohort. This substantially increases power and precision. However, it has the potential to alter participant profiles by mobilising those more averse to sanctions or by modifying motivations and thus individuals' actions (Redman 2020). This is also why testimonials and reinforcement of information about the optional nature of the programme were concurrently implemented. In any case, we managed to involve almost 4/10 single mothers in the encouragement groups.

3 Profiling the population of compliers

The average effect of encouragement on participation over the 4 cohorts is about 38.6 pp and we already saw that there is heterogeneity across cohorts. To understand the selection process, I start by estimating heterogeneity in the effect of encouragement on participation for a selected set of covariates: the same I use in all models. More precisely, I want to know if, *ceteris paribus*, the effect of encouragement is higher than the average effects for different groups of individual. To do that, I follow Lin (2013) and simply add de-meaned selected covariates and interactions with encouragement in the first stage. Denoting D for participation, Z for encouragement, I estimate the following regression using OLS (Lin 2013):

$$D_{ijc} = \alpha + \pi Z_{ijc} + \dot{\mathbf{X}}'\boldsymbol{\rho} + Z_{ijc}\dot{\mathbf{X}}'\boldsymbol{\beta} + \varepsilon_{ijc} \tag{2.1}$$

Where elements of $\dot{\mathbf{X}}$ are $\dot{X}_{ijc} = X_{ijc} - \bar{X}$ are the covariate deviations from the sample mean. Furthermore, I follow S. Athey and Imbens (2017a) and use only binary variables so this model is fully saturated and estimate the conditional expectation function perfectly. In this model, I don't use block fixed effect but add the blocking variable separately to estimate the average difference across variables. I weight observations by the inverse instrument-propensity score $\widehat{q}_{jc}$ estimated[34] using a probit of encouragement on block x cohort fixed effects to ensure conditional independence of the encouragement (Rosenbaum and Rubin 1983). Estimating equation (2.1) with weights gives the so called *doubly robust* estimator that ameliorates this sensitivity to misspecification (Hirano, Imbens, and Ridder 2003). Note that this model is akin to a Blinder-Oaxaca decomposition in which I interpret the *unexplained* part as treatment effect heterogeneity (See e.g. N. Fortin, Lemieux, and Firpo (2011) and Ding, Feller, and Miratrix (2019) for more details and alternative methods).

The covariates include cohorts, unemployment status, number of children, years receiving RSA, initial assessment, education, French citizenship, quartiles of age, receiving child support or family support allowance, housing benefits, children under 3, children 3 to 5, children older than 16, taxable income higher than median and quartile of baseline income per capita. Although it would be safer to let a machine-learning algorithm select the most relevant covariates, the dataset is small for a generalised random forest, and double-lasso may end-up creating omitted variable bias (Wüthrich and Zhu 2023). In this analysis, I chose simplicity, interpretability and consistency with other analyses although such analysis may be implemented in future revisions of the paper.

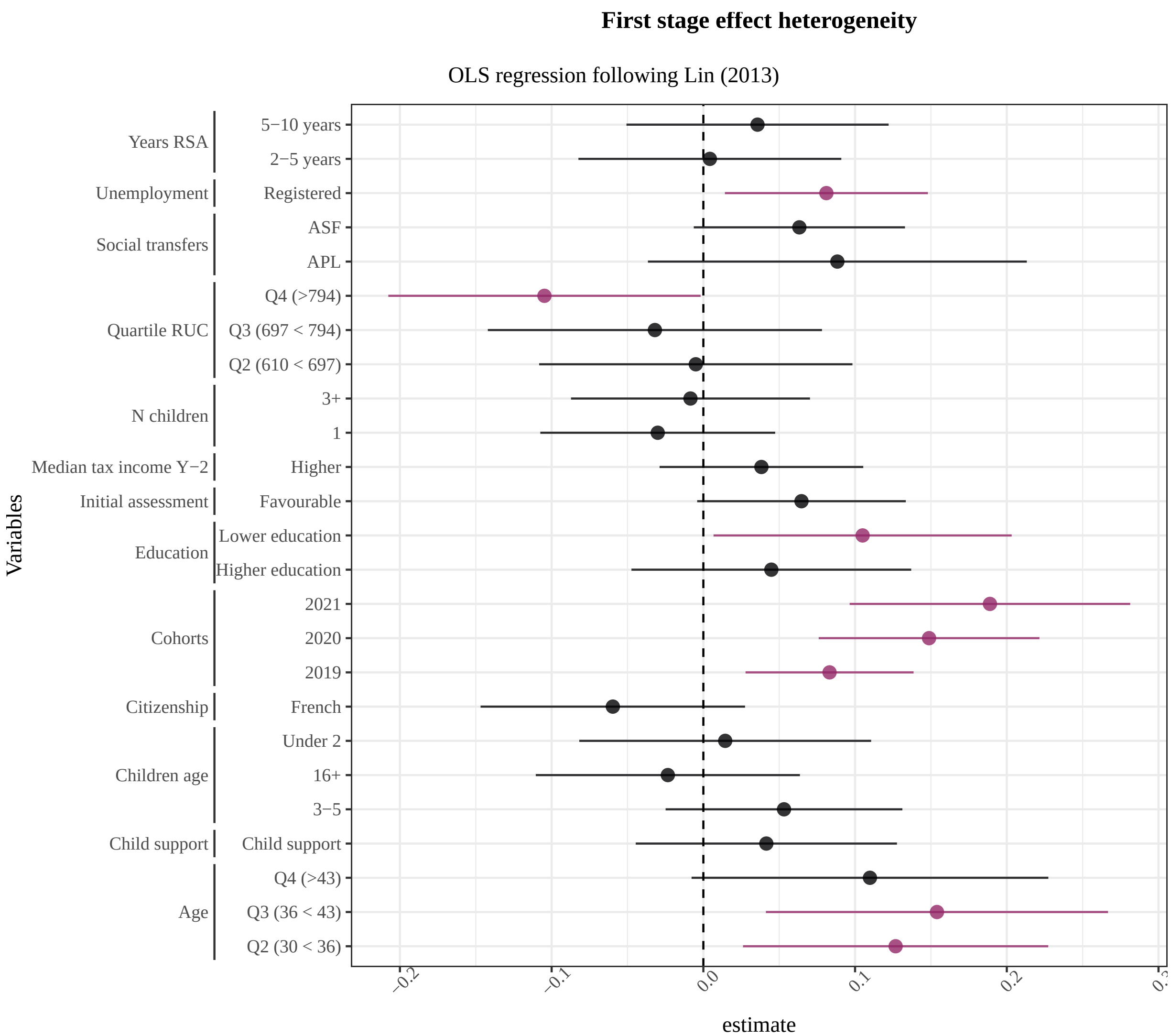
Higher participation among the least educated and poorest single mothers in their thirties Estimates of the $\widehat{\boldsymbol{\beta}}$ coefficients are presented in figure III.2 and confirm the previous observations on the heterogeneity w.r.t cohorts. The other results help defining a typical profile of compliers. First, those registered at the Employment agency and with favourable initial assessment are about 8 pp more likely to participate, with 95% pointwise confidence intervals between 1 and 15pp. Second, participation is higher for those with less than a high school degree (+ 10 pp) and the poorest at the time of random assignment. The 25 % with highest income are indeed -10pp less likely to participate than the 25% poorest. Age is also predicting higher participation for single mothers in their thirties, implying a lower participation among the youngest. Overall, these estimates shows that the programme attracted middle-aged mothers among the poorest and least educated of the sample. The larger take-up from those registered at the Employment agency may capture part of the latent *fear* of sanction or reflect a closer relationship with the labour market. In the control group, those registered at the Employment agency have a much higher job finding rate than unregistered mothers, as seen in Figure III.1. In the absence of intervention, registered single mothers are more

[34] The instrument-propensity score is simply the proportion of the encouragement group in each block and weights: $w_{jc} = \frac{Z_{ijc}}{\hat{q}_{jc}} + \frac{(1-Z_{ijc})}{1-\hat{q}_{jc}}$

likely to find a job, and their employment level after the period of the programme is the highest of all the compared groups.

Moreover, those who receive the family support allowance (ASF) and with favourable initial assessment are slightly more likely to participate although I cannot exclude 0 from the 95% CI. ASF is the allowance paid when the other parent is unable to pay child support or while waiting for a judge to set the amount. Those who do note receive it either receive child support from the other parent or did not go through the administrative process of setting child support (Périvier and Pucci 2019). The other plausible covariates I introduce do not predict participation. In particular, neither the number of year receiving RSA, nor the number of children predict participation although they predict outcomes.

Figure III.2: Estimating the heterogeneous effect of encouragement by baseline characteristics



Sources: ALLSTAT, observations one month before random assignment. Cohorts 2018 to 2021.
Notes: linear regression of participation on encouragement following Lin (2013): I regress participation on encouragement, demeaned covariates and the interaction of both. I use inverse probability weighting to account for the design.
I only present the coefficients of the interactions. The effect of encouragement and coefficients of the level variables are not shown (and very close to 0 by virtue of random assignment).
Point−wise 95 % confidence interval based on cluster−heteroskedasticity robust standard error with no degree of freedom adjustment (CR0).
Pink coefficients indicates coefficients that exclude 0 from the 95% CI.

4 Evolution of the main outcomes over the period

Figure III.3 illustrates the average employment levels over time, while Figure B.16 in the Appendix depicts the evolution of disposable income per consumption unit. Both figures include means with point-wise 95% confidence intervals and a spline smoothing by encouragement status and cohort. In addition, I plot the timing of the reform increasing in-work benefits in the first quarter of 2019 and the three lock-downs of the Covid-19 pandemic. These four panels offer a clear overview of the employment trends. While there are small variations between cohorts, there are no significant differences between encouragement and control groups after the end of the programme and no sign of improvement for earlier cohorts.

Before random assignment, we observe some Ashenfelter's dips (J. Heckman et al. 1998) with downward pre-trends and the lowest employment rate on the quarter before sampling. Cohorts 2018 and 2019 are balanced at baseline, but not cohort 2020 and 2021 for which differences between groups are reversed. However, these imbalances are constant for at least two years and are differenced-out in the estimations. The time span between random assignment and the beginning of the programme varies between cohorts, letting the control group get ahead of the encouragement group. This is particularly true for cohort 2019 and 2020, but not at all for 2018, whose recruiting period was the shortest. For the 2019 cohort, the PA reform occurred right in the middle the recruiting period, probably contributing to the observed lock-in effect.

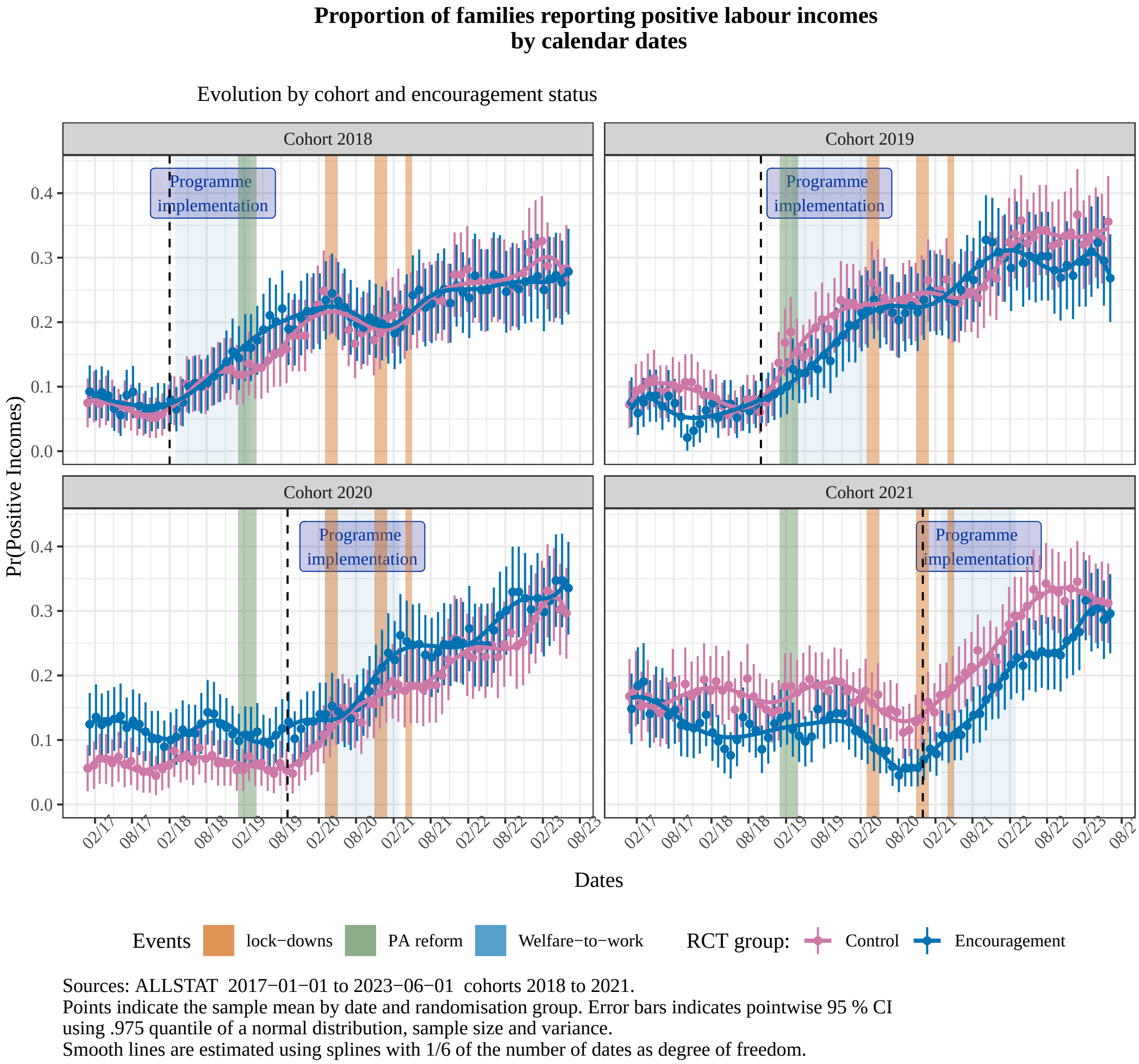
The first quarter of the training remains stable in all quarters while the qualitative analysis reports that this period is crucial in boosting motivation and self confidence, and start a positive dynamic. It, start showing up on the second quarter but the employment growth in the control group is still higher.

In the second half of the programme, the job finding rates of encouragement groups notably increase, even for the 2020 cohort in the midst of the pandemic. For the 2018 cohort, there is a short-lived positive difference favouring the encouragement group, which coincide with the reform of in-work benefits. But the control group caught-up 6 months later and employment rates are the same for the rest of the observation window. For the 2019 cohort, the programme ended right at the beginning of the pandemic with no difference between groups. Employment dynamics froze until the last lock-down where job finding rates started to increase again for both groups. If anything, the encouragement group has a slightly lower employment level than the control group. Finally, Cohort 2020 and 2021 ends-up with the same employment levels in both groups at the end of the programme. Note that ultimately, roughly 1/3 of the sample has positive labour income. This is true for all cohorts and randomised groups. The 2018 cohorts is the slowest to converge, and the pace seems to increase in recent cohorts despite the pandemic. Whatever the duration since we sampled the cohort, about 2/3 of each cohort remains without a job.

The dynamics of disposable incomes per consumption units presented in Figure B.16 in the Appendix is tightly linked to that of employment, although it shows meaningful differences. The encouragement group of the 2018 cohort is significantly poorer than the control group in the long run. The encouragement group of the 2019 cohort was significantly richer than the control group before random assignment and this difference correspond to the gap we observe at the end of the period. There are strong lock-in effects for all groups. The 2021 encouragement group was significantly poorer than the control group and the initial gap remains after the end of the programme.

Overall, these descriptive statistics show that on average, these households are and remain poor, with little evidence of an effect in intention-to-treat. I now describe my empirical strategy.

Figure III.3: Average proportion of families reporting positive work income over the period



IV Identification and estimation strategies

I mostly focus my analysis on four target parameters:

1) Average effects of encouragement on participation (First stage)
2) Average characteristics of compliers and never-takers
3) Dynamic and aggregated average intention-to-treat effects
4) Dynamic and aggregated average treatment effects on the treated.

Being the only experimental evaluation of such programme in France, I also estimate the dynamic average treatment effect on the treated in an event study design as if I had no randomised experiment. Comparing these estimates with those obtained with the instrumental variable strategy sheds light on the importance of selection bias.

Identification stems from random assignment of encouragement within blocks of single parents. At each point in time, there are control units of the same cohort and from the same block whose observed outcomes serve as counterfactual. Therefore, all well conducted comparisons between the encouragement and control groups have a causal interpretation. In particular, block-random assignment gives a causal interpretation to the first stage and comparisons of outcomes are interpreted as *intention-to-treat* parameters (S. Athey and Imbens 2017b). Then, assuming encouragement has no other effects on outcomes but through its effect on participation, an instrumental variable strategy yields the effects on compliers, which in this setting with one-sided compliance, can be interpreted as *treatment effect on the treated* (Frölich and Melly 2013). In practice, the main challenges are i) to aggregate treatment effects across blocks, cohorts and time without imposing restrictions on heterogeneity and ii) compute appropriate standard errors and correct for simultaneous inference.

1 First stage

It is well known that the average causal effect of the encouragement on participation across cohorts can simply be obtained by regressing participation on encouragement and block x cohort fixed effects using OLS and cluster-robust standard errors[35] (Negi and Wooldridge 2021; Alberto Abadie et al. 2022; C. de Chaisemartin and Ramirez-Cuellar 2022).

By the Frisch-Waugh-Lovell (FWL) theorem, this regression is equivalent to regressing 1) participation on block fixed effects 2) encouragement on block fixed effects and 3) residual of the first one on the second. The first auxiliary regression removes the average within blocks, therefore averaging control and encouragement together. Since there are no treated unit in the control group, their predicted values are out of the [0;1]. This model is therefore far from the conditional expectation function, as illustrated in Figure E.28 in the Appendix. While this misspecification is not a problem to estimate the average effect on participation, it has important consequences when the model serves as a first stage of a TSLS.

First stage for the second stage Since the projection matrix of the first stage is part of the solution of the TSLS system, the model is biased and does not recover the LATE. The recent work of Blandhol et al. (2022) revisits the use of TSLS to estimate the LATE and shows that the only specifications that have a LATE interpretation are i) "saturated" specifications that control for covariates non-parametrically or ii) models with restrictive parametric assumptions. Moreover, the fully saturated specification is more sensitive to finite-sample bias than estimates of just-identified models, especially in small blocks with low compliance[36]. Tübbicke (2023) compares different IV estimators and confirms that TSLS are likely biased when covariates predict the instrument and when groups are of different sizes.

In contrast, the recent work by Borusyak, Hull, and Jaravel (2022) provides an alternative but simple fix to this problem: de-mean the instrument by the block-specific instrument propensity score. The latter only depends on the proportion of encouraged households in each block that represents a *type* of household. The instrument propensity score is the assignment probability for each type and is given by the design, thus satisfying Assumption II of Borusyak, Hull, and Jaravel (2022). Let $\tilde{Z}_{ijc} = Z_{ijc} - \hat{q}_{ijc}$ be the "re-centred" offer instrument that subtracts the instrument propensity score q_{ij} from the encouragement indicator in block Z_{ij}. To estimate q_{ijc} I simply run a probit regression of Z on block-cohort fixed effects and use its predicted probability $\hat{q}_{ijc}$. I then estimate the following regression using OLS:

$$D_{ijc} = \sum_c \sum_j \alpha_{jc} A_{ijc} + \pi \tilde{Z}_{ijc} + \varepsilon_{ijc}, \quad i \subset \mathcal{P}_{\{\cdot, m=c\}} \tag{2.2}$$

Where D_{ijc} is a dummy for participation for parent i of block j from cohort c, A_{ijc} are block fixed effects (types) and $\tilde{Z}_{ijc}$ the demeaned instrument for encouragement. I estimate this model on the month of random assignment. To further improve the validity of this model, I weight observations by the inverse of the instrument propensity score (IPW) to obtain a so-called doubly-robust estimand. The weights used are $w_{IPW} = \frac{Z_{ijc}}{\hat{q}_{ijc}} + \frac{1-Z_{ijc}}{1-\hat{q}_{ijc}}$

[35] I discuss and justify the choice of clustering in a dedicated section below.

[36] The coefficient of the fully saturated regression are represented in Figure E.29 in the Appendix and the prediction in figure E.28.

To gain intuition, note that by the FWL Theorem, the estimand is unchanged if one includes q_{ijc} as control and the re-centred instrument $\tilde{Z}_{ijc}$i is replaced with the unadjusted offer Z_{ijc}. Since the propensity score is given by design, the main theorem of Rosenbaum and Rubin (1983) applies and this ensures unbiasedness of the IPW regression (Tan 2010; Matias D. Cattaneo 2010).

This model can accommodate additional covariates to improve precision, provided they are centred. Negi and Wooldridge (2021) show that there are no gain in a full regression adjustment i.e. interacting the treatment with demeaned covariates, when Z is balanced, which is the case here as shown in table B.6.

While these adjustment have little impacts on the estimation of the average effects as the propensity scores are close to .5 and only differ because of roundings in the assignment, they matter for the causal interpretation of TSLS, as I know discuss.

2 Characterising the full distribution of compliers' potential outcomes and attributes

Random assignment reveals potential participation and help identification of the effect on compliers (J. D. Angrist and Imbens 1995). To introduce notations for potential participation and other variables, let $V^k(d) = V^k(D(z)) \equiv V^k(d,z)$ denote the potential values of a variable V as function of participation and encouragement.

By design, $V^k(1,0) = \emptyset$ for there are no always takers. $V^k(1,1)$ is the revealed potential value of V for treated compliers and $V^k(0,0)$ the revealed potential value for the control group.

The key assumption for instrumental variable is the **exclusion restriction** where V is replaced by the set of outcomes Y:

Hypothesis 2.1 (Exclusion)**.**

$$Y^k(d,z) = Y^k(d,z') = Y(d)\ \forall\ z,\ z',\ d$$

which rule out an effect of encouragement, the instrument Z_{ijc}, on the variable of interest Y^k_{ijct} but through its effect on participation. In particular it means that for the never takers, $Y^k_{ijct}(D_{ijc}(0),0) = Y^k_{ijct}(D_{ijc}(0),1) = Y^k_{ijct}(D_{ijc}(0))$, which is actually the only restriction on potential outcomes we need (S. Athey and Imbens 2017b). In words, the encouragement has no effect on outcomes on its own, it only affects them through increased take-up in Reliance and the intention-to-treat effect is only driven by the treatment effect on compliers.

The exclusion restriction is a strong hypothesis and cannot formally be tested. In particular, it means that never-takers have no other reaction to the encouragement but to discard the invitation. Arguably, the encouragement might make people feel a little more scrutinised by the administration and threatened by welfare reduction if they do not comply. However the most effective way to avoid sanctions is to participate in the meeting which supports our hypothesis. Furthermore, this population receives letters from social services all the time and our invitation was no more threatening than any other. They are used to update their situation every quarters. I therefore assume the exclusion restriction holds.

The first stage allows to estimate the average proportion of compliers across cohorts $\pi = \Pr\left(D_{1ijc} > D_{0ijc}\right) = E\left(D_{ijc} \mid Z_{ijc} = 1\right) - E\left(D_{ijc} \mid Z_{ijc} = 0\right)$. The problem is that we don't know who the compliers are. When we see a particular observation among non-participants, it can be either never-taker or complier. But it can only be complier among participant because of one-sided non compliance. A well-known result discussed in A. Abadie (2003) and Frölich and Melly (2013) is that one can identify the distribution of the characteristics of the compliers. In practice, A. Abadie (2003) proposes a simple 2SLS procedure for characterising compliers, described by:

$$g\left(X_{ijc}, Y_{ijc}\right) \times \mathbb{1}(D_{ijc} = d) = \sum_c \sum_j \mathbf{A}'_{ijc}\alpha_{jc} + \gamma_d \mathbb{1}(D_{ijc} = d) + \mu_{ijc} \quad d \in \{0,1\} \tag{2.3a}$$

$$\mathbb{1}(D_{ijc} = d) = \sum_c \sum_j \mathbf{A}'_{ijc}\alpha_{jc} + \tilde{Z}_{ijc}\pi + \epsilon_{ic} \tag{2.3b}$$

where $g\left(X_{ijc}, Y_{ijc}\right)$ is any function of family baseline characteristics X_{ijc} and outcomes Y_{ijc}. Setting $d=0$ means I use Z_{ijc} to instrument $(1-D_{ijc})$ in an equation with $g\left(X_{ijc}, Y_{ijc}\right)(1-D_{ijc})$ as the outcome.

Borusyak, Hull, and Jaravel (2022) show how such specifications identify weighted averages of conditional-on-block IV coefficients[37]. In particular, the IV estimand is given by:

$$\gamma_d = \int w_d(t)\gamma_d(t)dF_j(t) \tag{2.4}$$

where $F_j(\cdot)$ gives the distribution[38] of blocks j and $\gamma_d(t)$ the block specific average of compliers.

This estimation thus recovers average characteristics of treated and untreated compliers:

$$\gamma_d = E\left[g\left(X_{ijc}, Y_{ijc}(d)\right) \mid D_{ijc}(1) > D_{ijc}(0)\right], \quad d \in \{0,1\} \tag{2.5}$$

By setting $g\left(X_i, Y_i(d)\right) = X_i$, the IV procedure produces the average of any predetermined covariate X_{ijc} for compliers within a block, and I estimate never-taker means by regressing $X_{ijc}(1-D_{ijc})Z_{ijc}$ on $(1-D_{ijc})\tilde{Z}_{ijc}$ with block fixed effect.

With this transformation, I can also estimate the distribution of the missing potential outcomes for compliers. By setting $g\left(X_i, Y_i(d)\right) = \mathbb{1}(Y_i \leq y)$ for a constant y and each value of d, I obtain the complier cumulative distribution functions of $Y_i(1)$ and $Y_i(0)$ evaluated at y.

To estimate the potential densities for treated and untreated compliers, Abdulkadiroğlu, Pathak, and Walters (2018) propose to use a symmetric kernel function $g\left(X_i, Y_i(d)\right) = \frac{1}{h}K(\frac{Y_i - y}{h})$ in equation (2.3a) to estimate the density of the potential outcomes for compliers at $Y=y$. $K(\cdot)$ is a symmetric kernel function maximised at zero and h is a bandwidth that shrinks to zero asymptotically. I follow Abdulkadiroğlu, Pathak, and Walters (2018) and evaluate complier densities at a grid of 100 points with kernel bandwidth defined as $h = 1.06 \times N^{\frac{-1}{5}}\sigma_d$, where N is the sample size and σ_d is the standard deviation of the potential outcome.

3 Dynamic treatment effects

I follow a design-based approach and consider two key parameters of interests for every relative month since random assignment:

- The average intention-to-treat effect of the programme
- The average treatment effect on the treated

Both are weighted average of block x cohort x time average effects. There are thus three key challenges. First, we need estimates of the block x cohort x time average effects, ITT and LATE. Second, we need to properly aggregate theme and third, we need to compute standard errors and account for simultaneous error.

With panel data, a natural way of estimating the dynamic treatment effects is to estimate cohort-time difference-in-differences and aggregate them in an event study. However, it is now well known that the staggered adoption of treatments with time-varying dynamic effects can cause sever issues in regression models without appropriate corrections or alternative estimands[39].

There are two design-based estimands that directly allow to estimate the event study parameters. My preferred models are based on the so called "stacked regression" which I present in details in the next sub-section. The other estimand is a modified version of Callaway and Sant'Anna (2020a) that requires separate estimates for each cohort and aggregation using the share of each cohort as weights. For intention-to-treat estimations, randomisation conditional on blocks ensure parallel trends and both methods yields causal estimates.

[37] See the chapter by J. Angrist, Hull, and Walters (2023) with a very clear presentation and application to school choice and the recent review by Borusyak, Berkeley, and Hull (2023) for more intuition on the theoretical results and review of other applications.
[38] Here, the distribution of weights is simply given by the share of each block in the sample.
[39] See e.g. J. Roth et al. (2021) for a recent review

A) Main model for dynamic intention-to-treat analysis:

The stacked-regressions consider each cohort as a sub-experiment and imposes a fixed event time windows relative to the month before random assignment. Let $m = t - c$ denote the relative time since randomisation such that $m \in \mathcal{M} =$ [-13,30], the upper limit being the last observation of the 2021 cohort. I estimate the following equations using OLS:

$$Y_{ijcm} = \sum_m \sum_c \sum_j \alpha_{jcm} A_{jc} + \sum_m \beta_m \tilde{Z}_{ijc} \times \mathbb{1}(t = c + m) + \sum_m \mathbf{X}' \boldsymbol{\rho}_\mathbf{m} + \varepsilon_{ijcm} \tag{2.6}$$

Where Y_{ijcm} is the outcome of household i in block j from cohort c, observed $m = t - c$ months since random assignment. A_{jc} are block $\times$ cohort fixed effects and $\tilde{Z}_{ijc}$ is the demeaned encouragement variable. I include a set of baseline covariates in a matrix $\mathbf{X}$ with a set of coefficient $\boldsymbol{\rho}_\mathbf{m}$ for each period (see below). Like before, I reweight observation by the inverse instrument propensity score to obtain doubly robust estimates. The interaction of all block $\times$ cohort with relative month dummies and the joint estimation of all ITTs yields a fully saturated regression. It estimates a treatment variance weighted average treatment effect of block specific intention-to-treat. One way to think about this regression is as a way of estimating all of the $ITT(c,t)$ parameters and then immediately aggregating them into a single set of event time parameters. Intuitively, it is equivalent to a regression of the long difference between month m and the reference month on the demeaned treatment variable and block cohort fixed effects restricted on the window of observations where cohorts are m months since random assignment. Each of these equations are stacked by interacting all right hand side parameters with relative time dummies. This method uses what looks like a typical TWFE regression estimate, but because of the structure of the data, it only incorporates clean controls. Notice that this model is similar to the event-study estimand of Sun and Abraham (2020). The only difference is that I use interactions of blocks instead of cohorts to better reflect the experimental design.

Adding covariates to the model The main motivation for adjusting for covariates is that the precision of the estimated average treatment effect can be improved if the covariates are sufficiently predictive of the outcome (Lin 2013). Moreover, I include one dummy for individuals resampled and one for those encouraged in the 2022 cohort to allow a specific trajectory for this subsample. Randomisation ensures there is no correlation in the population between Z_{ijc} and the covariates $\mathbf{X}_i$, which is sufficient for the lack of bias from including or excluding the covariates (S. Athey and Imbens 2017b). I follow Negi and Wooldridge (2021)'s recommendation and use pooled regression since $\mathbb{E}[p_{jc} = .5]$ in the random assignment case is precisely the condition that implies no efficiency gain from full RA even when there is arbitrary heterogeneity in the treatment effects. In particular, including baseline outcomes (centred) may increase precision while leaving the p-limit of the ITT coefficients unchanged. In practice, I add covariates interacted with relative time dummies to have a coefficient for each relative time period, allowing specific relative date effect of each covariate.

Inference Because I expect heterogeneous treatment effects across blocks, I use cluster-robust standard errors[40] (CR0 typically used in Stata) (Liang and Zeger 1986) at the block-cohort level following the recommendations of Alberto Abadie et al. (2022). It is worth noting that these corrections for clustering are asymptotically unbiased and converge to the *super population* parameter - the latter being rather undefined. A more design-based inference could be derived using results from Alberto Abadie et al. (2020). However, since we have both random sampling and random assignment, there are not much gains to expect and we can rely on the usual asymptotics. A last choice is about the degree of freedom adjustments and C. de Chaisemartin and Ramirez-Cuellar (2022) conveniently summarize their recommendations for practitioners. In all models, I do not adjust the degree of freedoms following their recommendations for models with fixed effects, clustering and blocks with less than 10 units.

Because the analysis focuses on the dynamics of the treatment effects at different points in time, the chances of type-II error increase and the test statistics based on point-estimate standard error will over-reject the null hypothesis of no treatment effect. Instead, I consider testing K null hypotheses[41] $H_0^1, \ldots, H_0^K$ individually and require that the family-wise error rate, *i.e.*, the probability of falsely rejecting at least one true null hypothesis, is bounded by the

[40] Cluster levels are equivalent to a regression on data collapsed by block $\times$ cohort averaging the full set of block specific treatment effects.

[41] Where K can refer to m or other sets of joint hypotheses.

nominal significance level $\alpha = .05$. In what follows I use adjusted p-values to describe the decision rules. Adjusted p-values are defined as the smallest significance level for which one still rejects an individual hypothesis H_0^j, given a particular multiple test procedure (Hothorn, Bretz, and Westfall 2008). By construction, I can reject an individual null hypothesis $H_0^k, k = 1, \ldots, K$, whenever the associated adjusted p-value is less than or equal to the pre-specified significance level α, i.e., $p_k \leq \alpha$. Most specifications consider the estimation of the ITTs over 12 months before randomisation and up to 30 months after with 4 cohorts or 45 months with 3 cohorts. Simultaneous inference for that many parameters costs a lot of power. To gain precision, I define $S \equiv S(m) = \lfloor m/6 \rfloor$ the relative semester since randomisation to average over 6 months and estimate average effects over each period.

The parsimonious aggregated models estimate the following equation:

$$Y_{ijcm} = \sum_m \sum_c \sum_j \alpha_{jcm} A_{jc} + \sum_s \beta_s S(m) \times \tilde{Z}_{ijc} + \varepsilon_{ijcm} \tag{2.7}$$

Like before, the Frisch-Waugh Lowell theorem makes clear what this equation identifies. The first part of the right hand side of the equation removes any average differences *between* blocks at every relative month and leaves the *within-block* variation. It also removes differences in the share of encouraged households across blocks - which is akin to subtracting the block-specific propensity score of encouragement from the treatment dummy. I come back to this point in the description of the instrumental variable estimation strategy. The estimations $\widehat{\beta}_s$ of this equations therefore delivers a variance weighted average of block-cohort-relative months specific treatment effects. As S. Athey and Imbens (2017b) note, in general, this parameter is not an unbiased estimand of the ITT over the permutation distribution of assignment.

B) Adaptation of Callaway and Sant'Anna (2020b)

Among the new methods for difference-in-differences with staggered adoption, the method proposed by Callaway and Sant'Anna (2020b) is very well fitted for this setting. Their event-study estimators:

1. Estimate a generalised propensity score based on the covariates
2. Estimate every 2x2 difference-in-differences between the reference point and another date between treated and untreated individuals in the cells with inverse-propensity score weightings similar to Alberto Abadie (2005). These are group-time treatment effects (GTTEs)
3. Aggregate the group-time GTTEs weighting them by the relative size of each group in the sample to compute the event-study coefficients.
4. Estimate standard errors accounting for multiple testing using wild cluster bootstrap.

The main difference with the previous method is a new estimation of the propensity score including covariates, the use of the doubly robust estimator by Sant'Anna and Zhao (2020) and simultaneous inference using a wild cluster bootstrap. The doubly robust estimator estimates the sample analog of the following equation:

$$\mathrm{ITT}(c,t) = \mathbb{E}\Bigg[\left(\frac{Z_{ic}}{\mathbb{E}[Z_{ic}]} - \frac{\frac{p_c((X))(1-Z_{ic})}{1-p_c((X))}}{\mathbb{E}[\frac{p_c((X))(1-Z_{ic})}{1-p_c((X))}]} \right) (Y_t - Y_{c-1}) \Bigg] \quad \text{under conditional parallel trend} \tag{2.8}$$

However, assumption 2 in Callaway and Sant'Anna (2020b) imposes that each unit i is randomly drawn from a large population of interest. Here, the sampling probability is the same within cohort, but varies across. That means that using the CS-2020 estimator on the full sample would use observations of the untreated units of other cohorts in the comparison group. A simple way to correct this is to run steps 1 and 2 separately for each cohort panels, and then aggregate the parameters in an event study using a similar framework. This is what I call the "modified CS-2020"[42]. This estimand is doubly robust and does not restrict heterogeneity.

[42] I thank Pedro Sant'Anna for his help in adapting the DID package. The R code for this function is available upon request

The aggregate estimates are obtained by weighting each group time treatment effects by the following weights:

$$\theta_{es}(m) = \sum_{c} \mathbf{1}\{t = c+m\} P(C = c \mid t = c+m) TT(c, c+m) \tag{2.9}$$

In words, each cohort-time average treatment effects is weighted by its relative size in the sample of observations at time $t = c+m$. Substituting Z by D in equation (2.8), I could obtain an estimate of the average treatment effect on the treated assuming conditional parallel trend with an appropriate control group. I estimate these models to analyse endogenous selection in section VI.

4 Treatment effects on the treated

To estimate the average treatment effect on the treated using the encouragement as an instrument for participation, I simply estimate TSLS system of stacked regressions.

$$Y_{ijm} = \sum_{m \neq -1} \sum_{c} \sum_{j} \mathbf{A}'_{ijm} \beta_{jcm} + \sum_{m \neq -1} \delta_m D_{ijc} \times \mathbb{1}(t = c+m) + \mu_{ic} \tag{2.10}$$

$$D_{ijc} \times \mathbb{1}(t = c+m) = \sum_{m \neq -1} \sum_{c} \sum_{j} \mathbf{A}'_{ijc} \alpha_{jcm} + \sum_{m \neq -1} \pi_m \tilde{Z}_{ijc} \times \mathbb{1}(t = c+m) + \epsilon_{ic} \tag{2.11}$$

Each block x cohort x relative time fixed effects instrument themselves in the second equation while participation at every month is instrumented by the interaction of the demeaned instrument with the corresponding relative month dummy.

Because of one-sided non compliance, these parameters correspond to the average treatment effect on the treated (Frölich and Melly 2013). This result on shift share IV immediately applies to block-random encouragement design[43].

Like the previous models, I adjust standard errors for clustering at the block $\times$ cohort level. In some settings, $m = f(t-c)$ can aggregate several months in a single parameter. The estimator is then a positively weighted average of the LATEs over the period (Mogstad, Torgovitsky, and Walters 2021). Just like for the ITT, the repetition of statistical tests increases the chance of false positive and I use the same method as with the ITT to correct for the family-wise error rate and use a 5% adjusted p-value to define 95% adjusted confidence intervals.

Remarks Equations (2.10) and (2.11) are actually an event study version of a *fuzzy difference-in-differences* (FDID) estimator that recovers the LATE if one further assumes a stable percentage of treated unit in the control group over time (C. De Chaisemartin and D'Haultfoeuille 2017, Theorem 1) which is ensured by design.

Adding covariates to equations (2.10) and (2.11) instrumenting themselves can improve precision but has two important consequences: first it imposes constant treatment effects and second, it loses the LATE interpretation and may introduce bias through the linear projection of the first stage that may be further away from the true conditional effect on participation. Adding covariates to a linear equation and interacting them with the treatment indicator seems like a natural way to account for non-random assignment of the instrument while allowing for heterogeneous treatment effects. Unfortunately, no result implies that this procedure generally uncovers the LATE (Słoczyński, Uysal, and Wooldridge 2022b). Nevertheless, this model with covariates can be informative. If we observe differences between the LATE (as estimated without covariates) and estimates assuming constant treatment effect and covariate-specific time effects, then the constant treatment effect assumption is likely false and motivates the analysis of heterogeneous treatment effect.

[43] This result has been applied in the school choice literature (Joshua D. Angrist et al. 2017; Abdulkadiroğlu, Pathak, and Walters 2018) and is central to my work on daycare assignment (Combe and Heim 2024). All the future work on evaluation lay on this set of recent results.

V Does it work ? Main results on labour and poverty

This section presents the main results of this evaluation: the effects of the programme on labour market participation and poverty. I first start with the dynamics of the treatment effects on labour market participation, disposable income per consumption units and risk of poverty. I also estimate the distribution of potential incomes for treated and untreated compliers to gauge distributional effects. I interpret these results and complete the picture with supplementary estimates in Appendix C. In brief, I find that the programme generates a strong lock-in effects that slows employment, reduces disposable incomes and slows exit from poverty. These negative effects fade-out by the end of the programme. In the post-treatment period, there are no average effects of the programme on employment and disposable incomes and 90 % of the sample remains in poverty. In the year following the programme, I cannot reject that the distribution of disposable incomes are the same.

1 Lock-in and no average post-treatment effects on employment or disposable incomes

Figure V.4 presents the main results on intention-to-treat effects on labour market participation aggregating the four cohorts. The top panel shows the average levels of employment in encouraged and control groups while the second panel shows the estimates of equation (2.6): the dynamic intention-to-treat effects, with or without covariates. I also present estimates using the modified version of the estimator of Callaway and Sant'Anna (2020a), estimating all treatment effects of equation (2.8) and aggregating them with weights of equation (2.9). Results are almost identical although the doubly robust estimator is both more precise and slightly more negative during the lock-in period.

Longer recruiting processes impose large opportunity costs to participants Before random assignment, both groups had a similar level of employment, which slowly decreased. This pattern is akin to an Ashenfelter dip (J. J. Heckman and Smith 1999) although there are no differences between groups. In the first half of the programme, job-finding rates are higher in the control groups while employment dynamics of the encouragement group only fasten after the first half of the programme. Lock-in effects are very common in the literature. During this period, job-search intensity may be lowered because there is less time to search for a job, and participants may want to complete an on-going skill-enhancing activity. A common and intuitive result often reported is the the longer the recruitment or the programme lasts, the larger the lock-in effect[44]. Over the years, we have extended the period between the draw and entry into the programme to maximise participation[45]. And while larger shares of the targeted groups signed-in, they also lost job opportunities waiting for the programme to start, or to benefit from it.

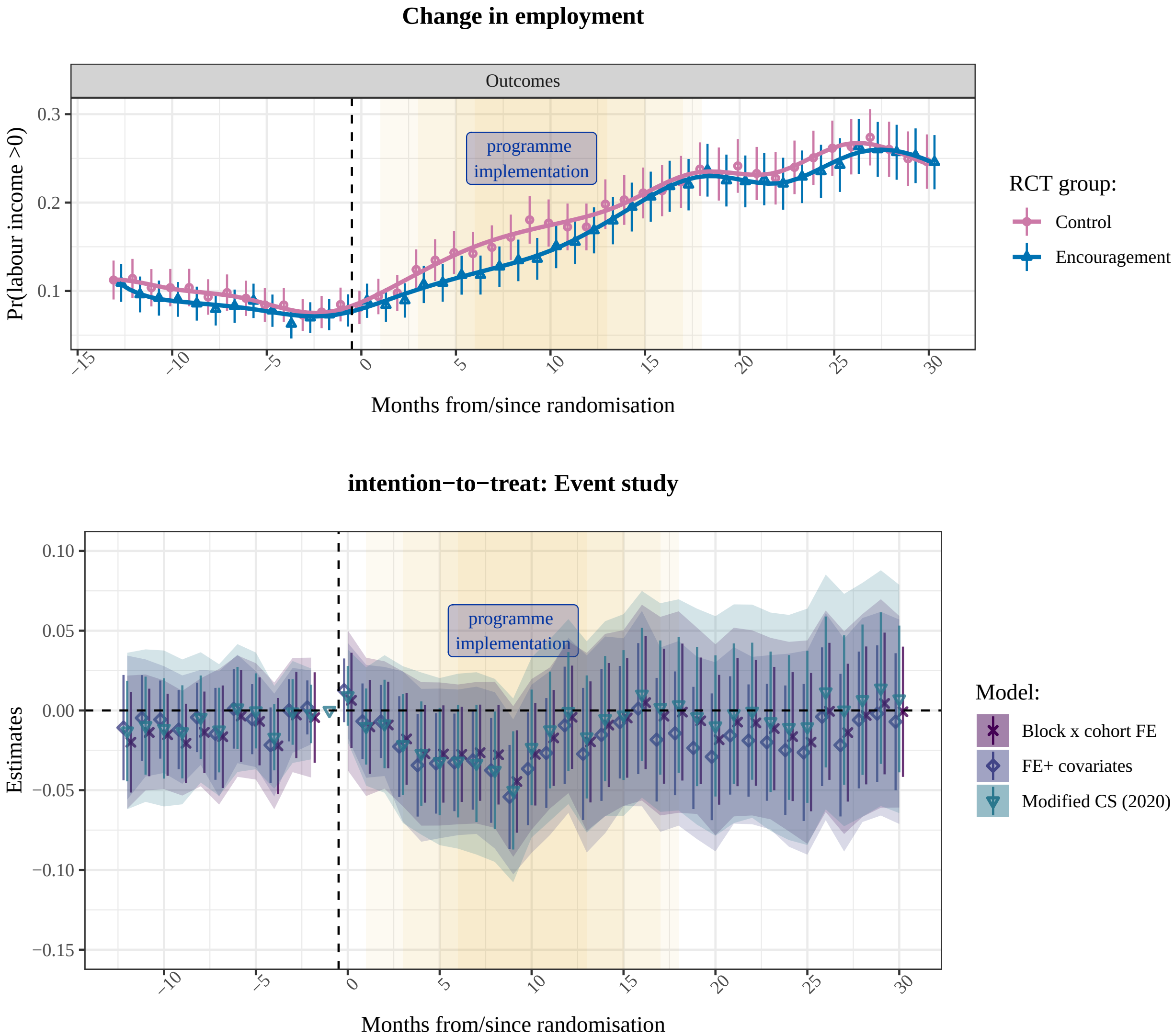
The qualitative report explains that the initial three months of the programme facilitated a rapid shift in momentum, boosting confidence and self-esteem relatively quickly. The group dynamic and "lever effect" of collective activities helped individuals overcome isolation. For the *slowed down* (mothers with access to resources but needing support to regain momentum), the pathway to stable employment or swift enrollment in training became viable. The *motivated* participants (mothers with limited social resources, often with migrant backgrounds, but determined and with some work experience) were poised for short-term employment opportunities, typically in sectors requiring minimal training and with higher recruitment rates.

These benefits show up on the job-finding rate after the 10^{th} month since randomisation. In one quarter, the encouragement group has gained on the the control group and after the programme, employment levels are virtually the same in both groups. Conversely, the qualitative evaluation also explains that isolated individuals with limited work experience and reduced autonomy (called the *excluded*), and those with active social networks but not actively seeking employment due to caregiving responsibilities, health issues, etc. (called the *hindered*) faced a longer timeline for employment reintegration, as their focus was not on immediate employment but rather on formulating a professional plan.

In the end, there are no average intention-to-treat effect after the end of the programme and the employment level is about 24.4% from 18 to 30 months after random assignment. In Figure E.32, I show the same estimations for

[44] See for instance Wunsch (2016) or Filges et al. (2015)
[45] An other reason is that it takes a lot of time to check every file before random assignment and managers did not want social workers overstaffed.

Figure V.4: Dynamic Effects of the programme on employment



Sources: ALLSTAT 2017−01−01 to 2023−06−01
Notes: The dependent variable is a dummy for positive individual labour income.
Top panel:
− Points indicate simple means by encouragement status over cohorts 2018 to 2021 in relative time since randomisation.
− Error bars indicate pointwise 95% CI using simple standard errors.
− Lines are conditional means estimated with spline regressions.
Bottom panel:
− Event study with block x cohort fixed effects with/without covariates using OLS or modified Callaway Sant' Anna (2020).
− Error bars, indicates 95 % CI using cluster−robust standard errors at the block x cohort level .
− Shades indicates 95%CI adjusting for the FWER uing the Holm−Bonferroni correction for OLS and wild cluster bootstrap for modified Callaway and Sant' Anna (2020).
Covariates are measured at the month before random assignment and include baseline level, initial assessment, education, French citizenship, quartiles of age, receiving child support or family support allowance, housing benefits, children under 3, children 3 to 5, children older than 16, taxable income higher than median and quartile of baseline income per capita.
All covariates are interacted with relative time dummies to have specific effects for each period.
For OLS models, covariates also include dummies for being resampled in the 2022 cohort and being encouraged, interacted with relative time dummies. For the modified CS model, observations of the late encouraged group are dropped.

the three first cohort up to 45 months. Results are less precise but qualitatively similar. They show no sign of improvement on the 15 additional months of observation.

Imprecise estimates still excluding policymakers's minimal expected effects Simultaneous confidence intervals barely exclude 0 on the 9^{th} month where the employment dynamics shift. Part of this imprecision and lack of statistical power - despite having one additional and larger cohort than initially - comes from unanticipated and large treatment effect variations across cohorts, and larger overall variance especially in the control group. The higher variability comes from difference macro-economic conditions while participants in the programme had rather similar trajectories[46].

To get a broader picture and estimate more precise parameters, I average the treatment effects over periods of 6 months and report the estimates in table C.7 in the Appendix. These estimates confirm the previous results and the-3.65pp of the ITT point estimate with covariates during the lock-in period correspond to -22.3 % of the employment level in the control group. Accounting for simultaneous inference, there are 95 % chance that the average ITT is between -7.3pp and 0 pp. The adjusted p-value for an effect different from 0 is 0.051. Using instrumental variables, The lock-in effect on employment for participant is -9.3 pp, with simultaneous 95% between -18.4pp and -0.3pp.

From 24 to 30 months after random assignment, the average treatment effect on the treated is -3.6pp, and I can rule out an average positive effect on the treated as high as 8.8 percentage point with 95 % confidence, thus excluding the minimal policymakers' target of 10 percentage points.

Lost opportunities, lost incomes Figure C.17 in the appendix completes the picture by presenting the average intention-to-treat effect on disposable incomes per consumption unit. This measure is directly computed by the National family allowance fund and accounts for family size. I use the Consumer Price Index for the bottom income quintile for actualisation in 2015 €. Obviously, the observed patterns are similar to the previous results but confidence intervals are more informative, and others are worth noticing. First, the seasonality in the top panel essentially stems from the imputation of one-time-per-year social transfers[47] by the Family allowance fund. Nevertheless, this panel shows that in real term, the average disposable income was and remained well below the poverty line and single mothers in this sample were getting poorer and poorer. The employment divergence noted above can be observed in the change in average outcomes. Between 6 and 12 months since random assignment, the control group had 727 $€_{2015}$ of income per capita and the encouragement group's was -30 $€_{2015}$ lower, which represent -4.1% of the control group average. The negative effect on disposable income takes longer to fade but in the end, both groups have the same average disposable income.

I present the parsimonious estimations over a 6-month period in Table V.2 and also include estimates of the average treatment effects on the treated (ATT) estimated and the average employment rate in the control group. The ATT is estimated instrumenting participation for each period by the demeaned encouragement interacted with period dummies and fixed effects. According to these estimates, participants lost an average of -76 $€_{2015}$ per month over the 6 first months of the programme. A back-of-the-envelop computation can give us the average income loss for compliers who did not get a job during the lock-in: If we divide this amount by the average treatment effect on employment discussed before (-9.3pp), we get 818. I discuss the effect of the programme on the intensive margin in Galitzine and Heim (2024).

This first set of results shows that the programme did not improve employment on average and caused participants to miss job opportunities during the lock-in period.

[46] See Figure III.3 in section III to see the outcomes by encouragement, cohort and calendar dates.

[47] Christmas and back-to-school allowances.

Table V.2: Aggregated effects of the programme on incomes per consumption units

	Mean control	**OLS**		**TSLS**	
		No covariates	*Covariates*	*No covariates*	*Covariates*
[-7 ; -1 [	709.7*** (7.4)	-5.2 (7.1)	-5.4 (6.0)	-13.4 (18.2)	-14.0 (15.3)
	[691.1, 728.3]	[-23.1, 12.7]	[-20.8, 10.0]	[-59.5, 32.8]	[-53.4, 25.5]
	adj.p.val. = 0.000	*adj.p.val. = 0.923*	*adj.p.val. = 0.878*	*adj.p.val. = 0.924*	*adj.p.val. = 0.878*
[0 ; 6 [	704.6*** (5.7)	-15.4** (6.9)	-16.1*** (4.2)	-39.5** (17.9)	-41.3*** (10.8)
	[690.1, 719.0]	[-33.0, 2.1]	[-26.9, -5.3]	[-84.8, 5.8]	[-69.1, -13.5]
	adj.p.val. = 0.000	*adj.p.val. = 0.107*	*adj.p.val. = 0.001*	*adj.p.val. = 0.111*	*adj.p.val. = 0.001*
[6 ; 12 [	727.3*** (7.5)	-24.5** (9.3)	-29.8*** (8.6)	-62.8*** (23.6)	-76.4*** (21.9)
	[708.3, 746.4]	[-48.1, -0.9]	[-52.0, -7.6]	[-122.6, -3.0]	[-133.0, -19.9]
	adj.p.val. = 0.000	*adj.p.val. = 0.038*	*adj.p.val. = 0.003*	*adj.p.val. = 0.036*	*adj.p.val. = 0.003*
[12 ; 18 [	727.3*** (6.9)	-14.7* (8.5)	-18.7** (7.8)	-37.6* (21.3)	-47.9** (19.5)
	[709.8, 744.9]	[-36.2, 6.9]	[-38.9, 1.5]	[-91.6, 16.5]	[-98.2, 2.4]
	adj.p.val. = 0.000	*adj.p.val. = 0.294*	*adj.p.val. = 0.083*	*adj.p.val. = 0.278*	*adj.p.val. = 0.069*
[18 ; 24 [	740.0*** (8.5)	-8.9 (8.4)	-13.1* (7.8)	-22.8 (21.0)	-33.3* (19.3)
	[718.6, 761.4]	[-30.1, 12.2]	[-33.1, 6.9]	[-75.9, 30.4]	[-83.1, 16.4]
	adj.p.val. = 0.000	*adj.p.val. = 0.727*	*adj.p.val. = 0.357*	*adj.p.val. = 0.721*	*adj.p.val. = 0.338*
[24 ; 30 [	755.2*** (10.2)	-10.0 (12.3)	-12.1 (11.6)	-25.5 (31.1)	-30.9 (29.0)
	[729.6, 780.9]	[-41.2, 21.1]	[-41.9, 17.6]	[-104.2, 53.1]	[-105.7, 44.0]
	adj.p.val. = 0.000	adj.p.val. = 0.884	adj.p.val. = 0.793	adj.p.val. = 0.883	adj.p.val. = 0.793
Num.Obs.	28700	57927	57927	57927	57927
R2	0.007	0.085	0.227	0.082	0.222
R2 Adj.	0.006	0.043	0.190	0.040	0.184
Covariates			X		X
Mean F-stat 1st stage				3274	2994

* p < 0.1, ** p < 0.05,*** p < 0.01 using point-wise p-value. Adjusted p-value and confidence intervals account for simultaneous inference using the Holm–Bonferroni correction. Standard errors are cluster-heteroskedasticity robust adjusted at the block x cohort level.

Notes: Control group means estimated using OLS with period dummies and no constant. OLS columns indicates average ITTs, TSLS columns indicate average ATTs. All models include block x cohort x relative time fixed effects and use inverse instrument propensity score weighting for double-robustness. Encouragement variable is centred by the instrument propensity score. I report the average of the F-stats for the first stages of all treatment periods.

Covariates are measured at the month before random assignment and include baseline level, initial assessment, education, French citizenship, quartiles of age, receiving child support or family support allowance, housing benefits, children under 3, children 3 to 5, children older than 16, taxable income higher than median and quartile of baseline income per capita. I also include dummies for being resampled in the 2022 cohort and being encouraged. All covariates are interacted with relative time dummies to have specific effects for each period.

No effect on the probability of welfare or in-work benefits but an increasing amounts of social transfers To complete this picture, I run additional estimations of the dynamic intention-to-treat effects on additional outcomes that were pre-registered: i) Receiving RSA (Figure C.19), ii) receiving the in work benefit (PA) (Figure C.20) and iii) total social transfers (Figure C.21). I find no intention-to-treat effects of the programme on RSA or PA. At best, it suggests an increase of RSA take-up at the onset of the programme showing social workers' work of making sure all participants had access to all social transfers they were eligible to. This statistically insignificant *bump* can also be seen on the ITT estimates on the probability to report quarterly incomes (testing differential attrition) in Figure B.13.

However, this programme, which was expected to reduce reliance on social transfers, actually causes an increase in the amount received. Figure C.21 shows a steady trend in the intention-to-treat effect on total amount of cash transfers, which excludes 0 from the simultaneous confidence interval of the very last period. The table C.8 in the Appendix presents the aggregated effects over 6-months periods and finds a significant treatment effect on the treated of approximately € 100. Looking forward in time with the first three cohorts, Figure E.34 shows similar estimates as those with 4 cohorts up to 30 months, but the effects keep on increasing, excluding 0 from the simultaneous confidence intervals for several months.

Note that participants do not get more social transfers than when they were recruited. They lose less over time. Moreover, this variable does not account for changes in family size. Dividing total cash transfers by the contemporaneous number of consumption units, the mediated effect on cash transfers disappears entirely, as shown in Figure E.35. Thus, the lesser decrease in cash transfers compared with the counterfactual stems from changes in family size, as captured by the number of consumption units.

2 Slow-moving poverty rates and little distributional effects

I measure poverty as a threshold defined by the poverty line in 2019: € 1040 current. This value is actualised like the disposable income in $€_{2015}$ using the Consumer Price Index for the bottom income quintile.

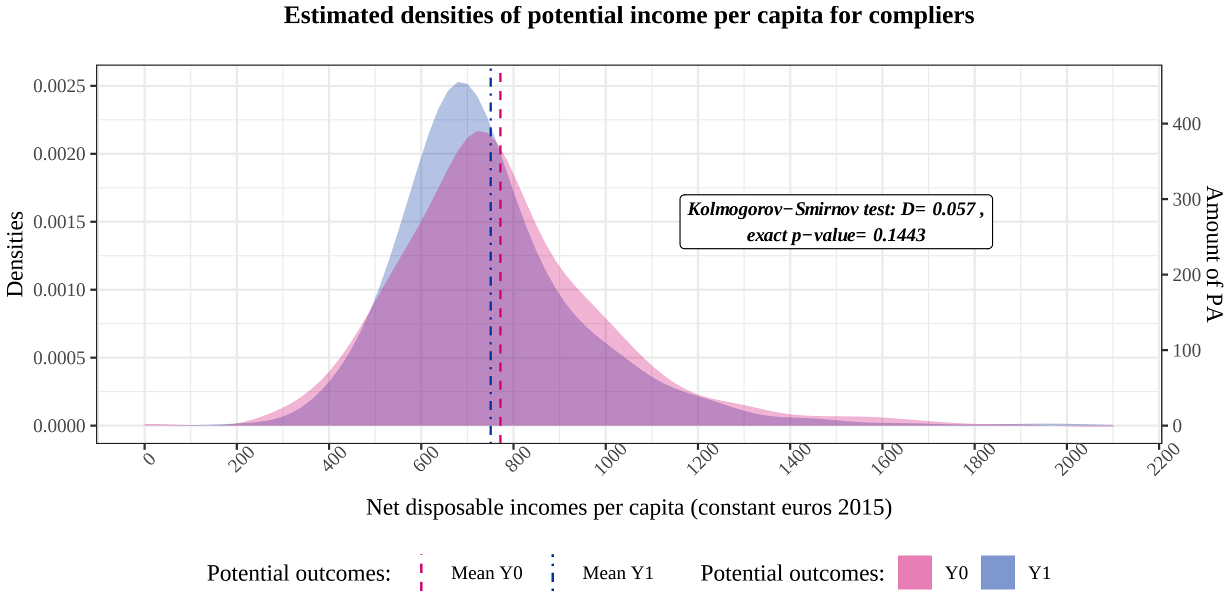
The programme slows the climb out of poverty Figure C.18 in Appendix C presents the average intention-to-treat of the risk of living with less than the poverty line. The top panel shows that at baseline, 98% of households live in poverty and, following the lock-in effect, we see the control group climbing out of poverty faster than the encouragement group in the first half of the programme, but a gap remains between the two groups that entirely close up at the very last period. The second panel estimates the intention-to-treat effects and shows a significant increase of poverty during the lock-in period. These estimates lack precision because of the number of parameters estimated. Aggregating the treatment effects over 6 months period and estimating the LATE, I find that the programme increases poverty incidence by 8 pp among compliers for the entire period of training, and exclude 0 from the adjusted 95% confidence interval (See Table C.9 in Appendix C).

Estimating participants potential incomes after the end of the programme To get a sense of the distributional effects of the programme, I estimate the counterfactual density of disposable incomes per consumption units from 18 to 30 months after the programme. To do that, I estimate equations (2.3b) and (2.3a) by 2SLS a 100 times on a grid of disposable incomes and using a kernel weighting of incomes as a dependent variable (See section IV for details).

The results on the years of observations after the end of the programme are presented in Figure V.5. My previous estimations of the average effects on participants were small and insignificant. This figure shows that the potential distribution of incomes for treated and untreated compliers are quite similar over the entire distribution. A Kolmogorov-Smirnov test comparing the individual average outcomes over the period between the encouragement and control group cannot reject that the two distributions are drawn from the same distribution with a p-value of 0.144. However, potential outcomes of compliers when they are treated (blue density) are slightly more concentrated around € 700 while the counter-factual density is a bit thicker between € 800 and € 1100.

These results contradict my initial hypotheses. At the onset, my priors were that this programme *could* make a change and help some of these vulnerable families get out of poverty through employment. Policymakers spent

Figure V.5: Distribution of the change in potential incomes per capita of compliers over 16 to 28 months (One year) after random assignment



Sources: ALLSTAT, restricted sample from 18 to 30 months and removing outliers (over 5 000 euros or the 99.7 percentile)
This figure plots marginal potential incomes per capita for compliers. Treated densities are estimated using 2SLS regressions of the interaction of a kernel density function and potential participation on the latter instrumented by encouragement.
Models only control for block x cohort x time to event dummies.
Untreated densities are estimated by replacing participation with one minus participation in this 2SLS procedure. All models use a Gaussian kernel and the Silverman (1986) rule of thumb bandwidth.
These estimates use a bandwidth of 64.3
The Kolmogorov−Smirnov test is run on a collapsed dataset of individual mean of the outcome over the period and compare the empirical CDF by encouragement group.
I test the sharp null that that incomes per capita are drawn from the same distribution using 99999 Monte carlo permutations.

four times as much as the usual budget for social support to fund a programme designed by practitioners, with a strong emphasis on *evidence based* social support. Yet, not a single assigned quantitative objective has been achieved. The programme seemed to have a positive impact on the first cohort but new data rejected this first initial findings. Analysis over a longer time frame for the first three cohorts show no improvement. However, the Covid-19 may have simply annihilate any benefit of the programme for the early cohorts. These conclusions will be confronted in future analysis including data from the fifth cohort and surveys.

It is worth noting that our results are in total contradiction with those of the qualitative evaluation. In the next section, I discuss plausible explanations, with a strong emphasis on endogenous selection.

VI Selection biais and heterogeneity

In this section, I investigate plausible mechanisms for the absence of effect on employment through two main points: self-selection and heterogeneity. First, I estimate difference in differences models comparing participants with control groups as if I had no randomised experiment. While the encouragement group has been randomly selected, participants did self-select into the treatment. The instrumental variable framework infer the share of compliers in the encouragement group and reweight the intention-to-treat effect by this proportion. If we had not randomise, we would not know the share of compliers and need an alternative identification strategy: typically matched difference-in-differences, In the economic literature, most researchers do not have the opportunity to run experiment of welfare to work programmes. There are far more quasi-experimental research typically using a difference-in-differences strategy to identify treatment effects. For instance, among the 39 studies included in the systematic review of ALMP for the unemployed of Filges et al. (2015), 25 used a timing-of-event and/or matching design while 14 were based on RCTs. If instead of comparing encouraged and control, I compare participants with another comparison group, the doubly-robust estimand of Callaway and Sant'Anna (2020a) could recover the average treatment effect on the treated under conditional parallel trend. This assumption is arguably much stronger than the instrumental variable framework which essentially rely on exclusion of alternative causal path between encouragement and outcomes but through participation. Here, a causal interpretation of these estimates requires that in the absence of treatment, participants would have had the same average outcomes than the comparison group, conditional on covariates. This assumption will be violated if there are unobserved time varying confounders. In contrast, estimations using the encouragement as an instrument and assuming exclusion hold true retrieve the average treatment effect on the treated. The comparison between the two estimates are therefore informative of the importance of selection bias.

1 Self-selection: Observing the unobservables

To analyse the endogenous selection, I estimate difference-in-differences models of the effect of the programme on employment comparing participants with i) never-takers, ii) all non participants, iii) the control group, iv) the control group matching participants on the set of selected covariates. Never takers are those of the encouragement group who refused the programme and suffer from selection bias most. The control group contains never-takers and compliers - although I don't know who is who. The estimates with control units are biased by never-takers. Matched observations from the control group is the best I can do without randomisation.

The programme attracts those with highest employment potential The top panel of the Figure VI.6 presents the average employment by actual participation, splitting the encouragement group between compliers and never-takers. It corresponds to what Susan Athey and Imbens (2017) call *as treated* or *per protocol* analysis. It shows that participants have a much higher employment rate after the programme ends than the average of the control group and the never-takers. The employment rate of the latters is much lower than that of the average control group and we have seen earlier that the weighted average of treated and never-takers equals that of the control group.

This simple plot gives a first mechanism for the absence of effect: the programme attracted the single mothers who were most likely to find a job whether they participated or not. This selection effect is very strong and suggests that most single mothers who *would* have work enrolled in the programme. The qualitative evaluation entirely missed the selection bias in their analysis and instead, saw the effects of the programme in the large share of participants working. The complementarity of both approach is precious to understand what happens *under the hood* and provide clear messages for policy makers.

In the bottom panel, I estimate difference-in-differences using the modified CS (2020) estimand discussed above and use different control groups. Since I do not compare participants by block, I use cluster robust standard errors adjusted at the household level to account for serial correlation. First, I estimate the model *as if I had no control group* comparing treated and never-takers. This model is the most heavily biased by endogenous selection and yields the highest estimates. Then, adding the control group reduces the estimated difference to 10 percentage points. Third, removing never-takers from the comparison group takes out a large share of the observed differences. Finally, flexibility controlling for the chosen set of observables with the doubly-robust estimand of Callaway and Sant'Anna (2020a) further reduced the estimates and confidence interval no longer exclude 0.

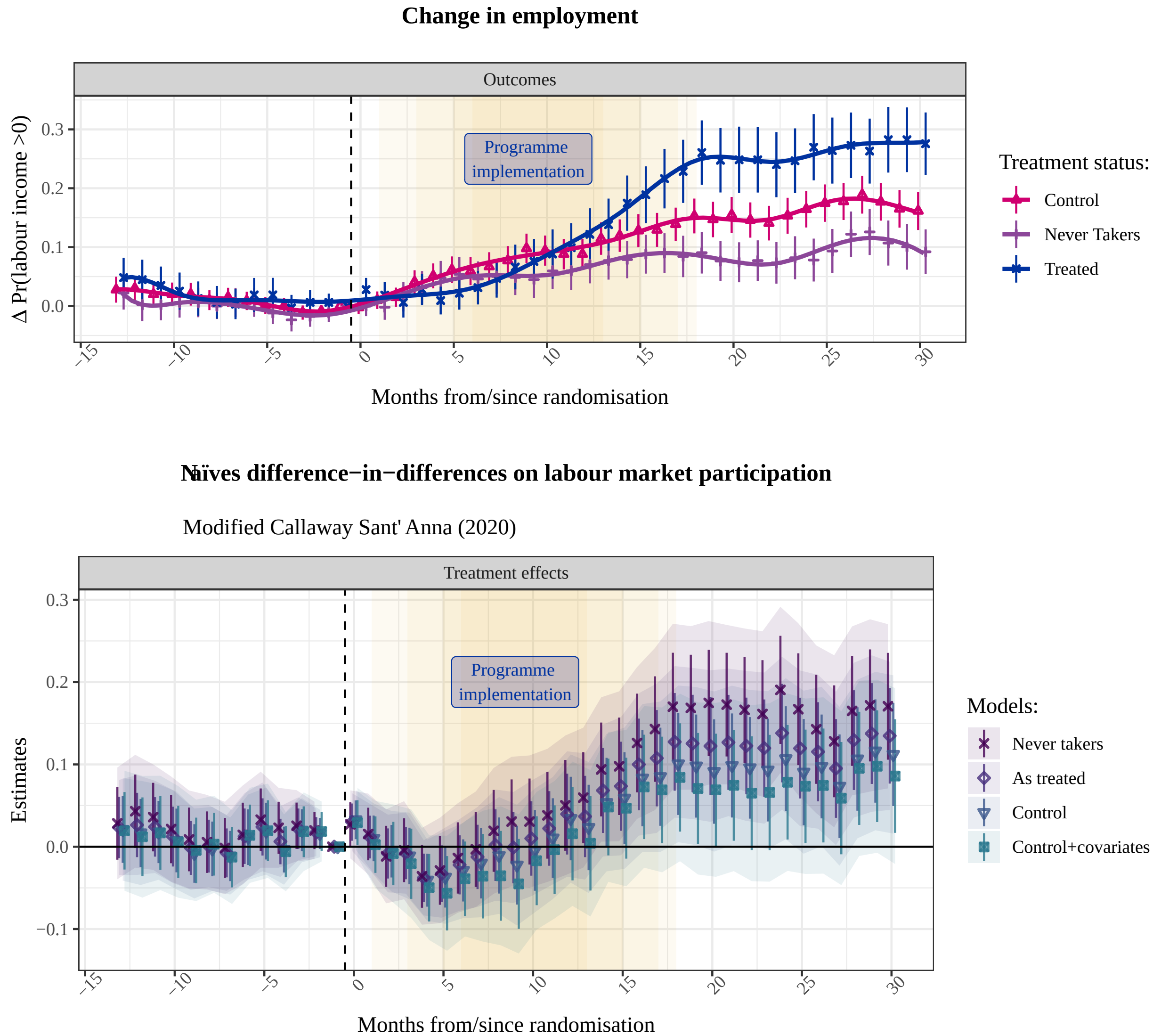
Visually, the event study is very convincing and the world is full of incentives to choose this model (Brodeur, Cook, and Heyes 2020). Pre-registration of the protocol, empirical strategy and main outcomes really made a difference. Over the past few years, it enabled me to dodge political pressures and critics on the choice set of outcomes. They also protected me from *HARKing*[48] my own research[49]. Difference-in-differences estimates are indeed very convincing. Estimates on lead months are close to 0 without trends, there is small lock-in and then a large increase in employment. However, all these results come from endogenous selection and collider bias due to the very nature of the target (Pearl 2009). The eligible population is conditioned upon receiving RSA for more than two years. Sampling is condition on past outcomes[50] and already partly controls for past employment spells. Moreover, the

[48] Hypothesis After Results Are Known, see Kerr (1998) for a discussion.

[49] Not by lack of scientific integrity but because these results using fancy, flexible, doubly robust new methods where more in line with my priors on the effectiveness of the programme. Researchers are not exempt from confirmation bias and neophilia, especially when it comes with new tools in lieu of old-school *boring* regressions.

[50] Which is not a problem *per se*, it is the sub-population of interest.

Figure VI.6: Comparisons of labour market participation by treatment status and difference-in-differences.



Sources: ALLSTAT 2017−01−01 to 2023−06−01 cohorts 2018 to 2021.
Notes: The dependent variable is the long difference between employment at any relative month and the month before randomisation.
Top panel:
− Points indicate simple means over cohorts 2018 to 2021 in relative time since randomisation with 95% error bars by treatment status.
− Lines are conditional means estimated with spline regressions.
Bottom panel:
Separate difference−in−differences for each cohort using the doubly robust estimator proposed by Sant' Anna and Zhao (2020) aggregated following Callaway−Sant' Anna (2020), i.e. separate DiD for each cohort and weighted average of cohort x time treatment effects.
Covariates are measured at the month before random assignment and include baseline level, number of years receiving RSA, number of children and unemployment registration status (uninteracted blocking variables). French citizenship, High/Low education, favourable assessment, receiving each social transfers, child support, children between 3 to 5 and at least one child over 16, quartiles of age, income per capita, taxable income.
The error bars indicate the 95% confidence intervals based on cluster−robust standard error adjusted at the household level. The shaded areas represent the FWER adjusted 95% confidence levels estimated using wild cluster bootstrap.

variables used to define blocks are not fixed attributes. For instance, the blocks of registered unemployed parents at the time of randomisation can be composed of long-term unemployed and very recent registrations. Without the experiment, this variable would be considered *endogenous* and require its own instrument. Said differently, eligibility varies with time and attributes measured at $t = c$ can be affected by past outcomes. In this setting, testing pre-trend can be very misleading. This is why I talk about *attributes* rather than characteristics: covariates are just observables at some points in time, held fixed and used to remove variation correlated with these attributes. Once we simply follow households, there are no other control and time varying differences between participants and others can affect the outcomes.

Difference-in-differences would reject the experimental results with 95% confidence. Table VI.3 displays the difference-in-differences estimates aggregated over months using long difference in employment as dependent variable, replacing the encouragement by participation in equation (2.7) and using only the control group for comparison (hence removing never-takers from the sample). The model with minimal covariates only includes specific trend by baseline outcome level. Models with covariate include the same set used in every estimations (see the figure notes). I compare these estimates with the ATT estimated by instrumental variables using long difference in employment as the outcomes. I do not control for baseline level in the first IV model to have no restriction on the treatment effect heterogeneity. Adding covariates does imply constant treatment effects but increase precision. I also present the estimate of the average *missing* compliers' potential outcomes $Y(0)$ following A. Abadie (2003). It is obtained using $(1-D_{ijc})Y_{ijcm}$ as dependent variable and $(1-D_{ijc}) \times S(m)$ instrumented by $\tilde{Z}_{ijc} \times S(m)$, the demeaned instrument for the relevant period. It shows that in the lock-in period, an average of 20% of participants would have worked had they not signed in the programme, although the average of the treated group is about 13%. Comparing estimations between instrumental variables and difference-in-differences, the estimated coefficients are generally of opposite sign and the 95% confidence intervals for the difference-in-differences estimates controlling for Family-wise error rates exclude the experimental point estimates from the lock-in period to the end. In other words, the best one could do with non-experimental methods cannot recover consistent estimates. Although not surprising (Lalonde 1986; J. J. Heckman, Lalonde, and Smith 1999; Card, Kluve, and Weber 2018), these results are worrisome if one thinks of the share of published research using such strategies to estimate the effect of similar programmes. I come back to this point in the discussion.

Table VI.3: Aggregated treatment effects on the treated on labour market participation

	Compliers' Y(0)	**TSLS**		**DID**	
		No covariates	*Covariates*	*minimal covariates*	*All covariates*
[-7 ; -1 [	0.099*** (0.031)	-0.024 (0.034)	-0.017 (0.023)	0.003 (0.010)	0.003 (0.010)
	[0.022, 0.175]	[-0.109, 0.061]	[-0.076, 0.042]	[-0.023, 0.029]	[-0.022, 0.029]
	adj.p.val. = 0.006	*adj.p.val. = 0.927*	*adj.p.val. = 0.944*	*adj.p.val. = 1.000*	*adj.p.val. = 0.999*
[0 ; 6 [	0.120*** (0.027)	-0.036 (0.030)	-0.039 (0.026)	-0.016 (0.013)	-0.018 (0.013)
	[0.053, 0.188]	[-0.113, 0.040]	[-0.106, 0.028]	[-0.051, 0.018]	[-0.052, 0.016]
	adj.p.val. = 0.000	*adj.p.val. = 0.627*	*adj.p.val. = 0.473*	*adj.p.val. = 0.693*	*adj.p.val. = 0.577*
[6 ; 12 [	0.208*** (0.036)	-0.073** (0.034)	-0.093*** (0.035)	-0.022 (0.016)	-0.025 (0.017)
	[0.117, 0.298]	[-0.159, 0.013]	[-0.184, -0.003]	[-0.064, 0.020]	[-0.069, 0.019]
	adj.p.val. = 0.000	*adj.p.val. = 0.128*	*adj.p.val. = 0.039*	*adj.p.val. = 0.607*	*adj.p.val. = 0.523*
[12 ; 18 [	0.261*** (0.045)	-0.016 (0.042)	-0.032 (0.042)	0.044* (0.023)	0.038* (0.022)
	[0.149, 0.374]	[-0.121, 0.090]	[-0.139, 0.074]	[-0.015, 0.102]	[-0.019, 0.096]
	adj.p.val. = 0.000	*adj.p.val. = 0.997*	*adj.p.val. = 0.927*	*adj.p.val. = 0.241*	*adj.p.val. = 0.355*
[18 ; 24 [	0.341*** (0.047)	-0.022 (0.047)	-0.052 (0.042)	0.076*** (0.025)	0.071*** (0.025)
	[0.224, 0.458]	[-0.140, 0.096]	[-0.160, 0.056]	[0.012, 0.140]	[0.007, 0.135]
	adj.p.val. = 0.000	*adj.p.val. = 0.989*	*adj.p.val. = 0.665*	*adj.p.val. = 0.012*	*adj.p.val. = 0.023*
[24 ; 30 [	0.365*** (0.047)	-0.021 (0.049)	-0.036 (0.048)	0.077*** (0.028)	0.072** (0.029)
	[0.247, 0.482]	[-0.146, 0.104]	[-0.160, 0.088]	[0.004, 0.151]	[-0.003, 0.147]
	adj.p.val. = 0.000	adj.p.val. = 0.994	adj.p.val. = 0.940	adj.p.val. = 0.034	adj.p.val. = 0.069
Num.Obs.	56749	56749	56749	39344	39344
R2	0.120	0.106	0.274	0.240	0.266
R2 Adj.	0.079	0.064	0.239	0.188	0.212
Baseline level			X	X	X
Covariates			X		X
Mean F-stat 1st stage	3198	3223	2928		

* p < 0.1, ** p < 0.05,*** p < 0.01 using point-wise p-value. Adjusted p-value and confidence intervals account for simultaneous inference using the Holm–Bonferroni correction. Standard errors are cluster-heteroskedasticity robust adjusted at the block x cohort level.

Notes: the dependent variable equals 1 when the parent has positive labour incomes, 0 otherwise. Compliers' average obtained by TSLS of DY on D instrumented by centered instrument. Second and third columns report TSLS estimates of the ATTs with or without adding covariates in the model. I report the average of the F-stats for the first stages of all treatment periods. Columns DID remove nevertakers from the sample and compare treated and controls in difference-in-differences. Covariates are measured at the month before randomisation and include baseline level, initial assessment, education, French citizenship, quartiles of age, receiving child support or family support allowance, housing benefits, children under 3, children 3 to 5, children older than 16, taxable income higher than median and quartile of baseline income per capita. Observations that are treated in the cohort 2022 are dropped once they are resampled.

2 Puzzling heterogeneous treatment effects by children at baseline

In Appendix D, I provide estimates of the average treatment effect on the treated for employment and disposable income per consumption unit, categorized by various subgroups determined by blocking variables. Two additional exploratory analyses, not pre-registered, are also presented. All models incorporate covariates to enhance precision and are estimated over 6-month periods using Two-Stage Least Squares (TSLS) with demeaned instruments and block fixed effects.

Heterogeneous effects by cohort Figure D.22 illustrates the heterogeneous effects by cohort. Notably, the 2018 cohort exhibits negligible impacts on both employment and incomes throughout the entire period. Conversely, the 2019 cohort experiences a pronounced lock-in effect on employment, driven by a temporarily higher job finding rate for the control group post-PA reform. This effect diminishes by the end of the training, with no significant impact on disposable incomes. The 2020 cohort shows minimal effects on employment but a significant lock-in effect on incomes, dissipating rapidly. In contrast, the 2021 cohort reveals a distinct pattern, with negative employment effects intensifying post-lock-in, implying a -15 pp lower probability of employment 2 years after randomisation. While no employment effects are observed in the last 6 months, disposable incomes follow a similar trajectory, with a significant increase over the initial 18 months. There are no significant effect or heterogeneity on disposable income after 2 years since random assignment.

Treatment effects by number of children at baseline Figure D.23 presents the ATTs by the number of children at baseline. Participants with one child experience a significant, lasting negative effect on employment, peaking at about 17 pp by the programme's end. This group also encounters a substantial lock-in effect on disposable incomes, approximately €150. However, the income effect dissipates post-programme, with no lasting impact. Conversely, those with three or more children exhibit the opposite pattern, though estimates lack significance. They experience no lock-in effect on employment or income, with a positive employment effect stabilising around 10 pp higher (CI between -5 and 35 pp). The effect on disposable incomes is increasingly negative but not statistically significant.

Treatment effects by other subgroups Comparisons by the number of years receiving RSA and registration with the Employment agnecy (Figure D.24 and D.26 respectively) reveal minor differences, with slightly worse effects for those registered at the Employment agnecy and those with fewer than 5 years receiving RSA.

Figure D.25 illustrates outcomes by social workers' initial assessments, revealing more negative effects among those with a favourable assessment, while others experience increasingly negative effects on income per consumption unit with little to no impact on employment.

Finally, estimations by income levels at baseline presented in Figure D.27 indicate substantial negative and lasting effects on employment for those with incomes below the median. Interestingly, this group, more likely to participate, sees no effect on incomes. In contrast, those with higher baseline incomes exhibit no employment effects but face a significant decrease in disposable incomes by about €100 per month.

In summary, while the programme demonstrates little to no average effect on employment and disposable incomes, the heterogeneous analyses unveil noteworthy patterns, particularly regarding the number of children at baseline. Altogether, these results suggest heterogeneous changes in the composition of incomes which constitute a new puzzle that I address in Galitzine and Heim (2024).

VII Discussions

In this paper, I analysed the effects on labour market participation and poverty of an intensive welfare-to-work programme for single parents on long-term welfare in France. This experiment relied on a staggered block-randomised encouragement design to recruit participants from 2018 to 2022, building a dataset of 2073 households from 5 cohorts. I follow the 4 first ones (1671 households) with panel data from administrative records up to 30 months after random assignment. The main findings can be summed-up as followed:

1) The take-up is 38% on average, increasing from 28% to 48% from the firth to fourth cohort. This improvement most likely stems from adjustments to the recruiting process including i) more threatening invitations ii) information sessions moved to the programme's premises iii) change from collective to individual face-to-face meetings with project managers and iv) testimonies from former participants.
2) Compliers are more likely to be in their thirties, among the poorest, have less than a high school diploma, and be registered to the Employment agency.
3) The programme slows the job finding rate during its first half, causing a strong lock-in effect showing up on poverty rate, disposable income and employment. These anticipated negative effects fade-out by the end of the programme and there are no average effect in the post-treatment period.
4) The programme gradually increases the amount of cash transfers received, although this effect is entirely mediated by increasing family size.
5) Analysis of the heterogeneity of treatment effects reveals puzzling patterns, especially among high/low income and number of children at baseline, where effects on disposable incomes and labour market participation have diverging patterns, suggesting changes in the composition of incomes, corroborated by the increase in cash transfers at the end of the period.
6) The programme attracts those with highest potential employment levels, but do not increase labour market participation. The selection bias is so strong that estimates using the next-best identification strategy implemented with modern robust estimands fail to include the experimental results in the confidence intervals. Without random assignment, one would have wrongly concludes that the programme increases employment.

The effects of this programme on labour market participation and disposable income of poor single mothers in France are very negative. It failed at increasing employment or get these single parents out of poverty. In this section, I discuss plausible interpretations, threats to their validity and some policy implications.

A diversity of profiles with heterogeneous reactions In this experiment, compliers seem to be divided into two groups: the poorest with low education on the one hand, and those closer to the labour market on the other hand. For the formers, the programme reduces employment but has no effect on disposable incomes while for the latter, it has no effect on employment but significantly reduces their incomes. This can only mean that the poorest found other sources of incomes while those closest to the labour market earn less than they would have had they not participated. The total transfers by Family allowance funds are higher due to higher number of consumption units. This means that the programme affected family size and composition, possibly increasing cohabitation, fertility or through lower custody loss. Understanding these complex reactions are out of the scope of this study. In Galitzine and Heim (2024), we analyse these mechanisms and link compliers reactions to the incentives of the tax-benefit system, which, as it turns out, has a 10-15% implicit marginal tax rate at part-time minimum wage, and a 70-75% at the full time minimum wage. Estimating potential labour incomes distribution, we show that treated compliers bunch at part-time minimum wage and find a hole in the distribution for incomes higher than minimum wage. The qualitative evaluation also reports that among participants of the 2019 cohort on permanent or fixed-term contracts, 72% work part-time (58% on fixed-term part-time contracts) and 75% say they choose to do so[51]. Moreover, incentives are different by number of children and cohabitation with a partner, and we also find consistent treatment effects on the probability of re-partnering and de-cohabitation with older children. However, we confirm the absence of treatment effects on disposable incomes over the entire distribution, no matter how many children they had at baseline. Together, these two studies show that the programme helped single mothers optimise their circumstances, reducing work hours and incomes for those who would have worked, while also adjusting family size and composition, such that the

[51] Source from the internal data from the Departmental council. 5% say they they did not choose it and 20% did not answer.

resulting higher social transfers compensate lower labour incomes, maintaining the same standards of living as in the counterfactual. This may also include income effects when social workers helped participants access local aids or discounts - things I don't observe in the data.

Mothers and children in poverty and job quality Although the programme helped participants optimise, 89% remains in poverty 30 months after random assignment and 66 % have no job. We supported 328 single parents with their 672 children and 30 months after random assignment, 277 participants and their 568 children remain poor. Even among those who work, poverty remains strikingly high with 75% working poors among participants 30 months after randomisation. Despite social transfers, low-paid jobs do not provide sufficient income to support families and often do not meet parents' constraints (P. M. Evans 2007 ; Jaehrling, Kalina, and Mesaros 2015; Millar et al. 2018; Van Winkle and Struffolino 2018). Moreover, they do not allow mothers to build human capital (Blundell et al. 2016) and prevent them from investing in their children' education (Løken, Lommerud, and Holm Reiso 2018). Moreover, exposure to poverty during childhood significantly impacts various aspects of psychological well-being trajectories and learning abilities, with long term consequences (Lucas et al. 2008; G. J. Duncan, Ziol-Guest, and Kalil 2010; G. J. Duncan et al. 2012b; G. W. Evans and De France 2022). Conversely, activation with high social transfers has been shown to durably get single parents out of poverty (Markussen and Røed 2016).

In the qualitative study, some participants felt that their aspirations were not sufficiently considered. This includes mothers who were directed towards jobs or sectors that did not interest them or felt "pushed" into employment despite their difficulties. In fact, a large share of offers were jobs in the personal care sector (home help and care assistants), in supermarkets (checkouts or stock replenishment), or in cleaning services. Some individuals were not ready or equipped to face the working conditions proposed in low-skilled sectors and/or the time frame set by Reliance. In some cases, the person's initial project may have been out of sync with their current skills and/or the functioning of the job market, leading to a redirection towards job offers deemed more "realistic" by social workers. This paternalistic view is well documented and tends to create a sense of loss of control in people's life. Campbell et al. (2016b) reviewed qualitative studies on welfare-to-work programmes for single mothers, finding that programmes' demands often clash with parenting duties, resulting in precarious and low-paying jobs. Health issues like stress and depression are common, but some participants experience increased self-esteem. However, these programmes can reduce control over employment and childcare, contributing to health concerns. Similar findings were noted in the literature review of Baronnet et al. (2021).

Unresolved issues and possible hidden effects For others, Reliance's support helped realise that employment is not their short or medium-term goal. In most situations, participants are still overwhelmed by issues that strongly impact their daily lives, preventing them from envisioning a return to employment: lack of personal housing, child placement, serious illness, a child's disability, etc. The persistence of these trajectories reveals the lack of insurance against the consequences of certain life shocks. These situations, described by Perrin-Heredia (2009), highlight that for the most deprived social categories, "*life accidents are not random. [They] are both more likely to experience these accidents [...] and have fewer means to face them.*".

However, they may still benefit from the quality of social support, but not through employment. Indeed, the qualitative evaluation reports high satisfaction, increased motivation and self esteem, while reducing loneliness and isolation. PTSD and depressions were common among participants, some resulting from intimate partner violence and coercion (IPVC). Other research showed that such programme can have positive effects on IPVC-related PTSD (Meisel, Chandler, and Rienzi 2003; Perez, Johnson, and Wright 2012; Latzman et al. 2019).

The development of models to measure the effects of welfare programmes began in the late 1960s and early 1970s with the static labour supply model and was well worked out by the 1980s, extending to education, marriage, fertility savings and so on. And yet, as Chan and Moffitt (2018) point out, "*there has been far too little work on the dynamic aspects of labour supply choices in the presence of different kinds of programs (traditional welfare versus earnings subsidies, for example) where human capital, family structure, migration, occupational choice, and other lifecycle decisions are important*". Economics may well have missed important phenomena dismissing these dimensions. At this point, the lack of measures on other outcomes may put too much weight on the negative results of this evaluation. The analysis of surveys left for future work will help gauge the effects of the programme on subjective health and well-being.

The recruiting process was successful Targeting vulnerable populations often produces low take-up, which is attributed to various factors such as low monetary or utility gains, stigma associated with social programmes, monetary, non-monetary and opportunity costs of participation, imperfect information, administrative barriers, and inadequate eligibility measurements (Friedrichsen, König, and Schmacker 2018 ; Ko and Moffitt 2022). In this experiment, the lowest take-up is 8 points higher than a similar experiment from 2006 in Seine-Saint-Denis, the poorest suburb close to Paris (Crepon et al. 2013). After four years of implementation, we managed to enroll almost half of the 2021 cohort. For that, we played on both *threats* and *pull* levers in the recruiting process. First, we added an ambiguous yet threatening sentence on *"rights and duties"* in the invitation letter: a *threat effect* to foster participation. Second, we made recruiting sessions more warm, welcoming and individualised. We moved from collective information sessions in the *welfare-to-work* division of the Departmental council to individual face-to-face interviews with project managers in the newly renovated and well equipped premises of the programme. Additionally, former participants were involved in the recruiting sessions to answer questions and provide (positive) testimonies as peers. Higher take-up likely stems from these adjustments as corroborated by the qualitative evaluation. However, the Covid-19 pandemic and resulting restrictions of social contacts may have played a role as well, increasing the need to break isolation and loneliness. Furthermore, random sampling and assignment are very different from usual referral practices and may have reached households who would not have been proposed such programme.

Researchers and policy-makers have implemented various *carrots vs. sticks* schemes to foster participation in ALMP. Recent experimental works show that receiving information is not enough ; transaction costs and administrative barriers matters and human contacts help foster participation. For instance, Chareyron, Gray, and L'Horty (2018) evaluate how changing the content of the official letter sent to RSA recipients could foster participation in welfare-to-work programmes. Randomising wordings of different letter templates, they find no effect of either the "simplified letter" or the one highlighting the potential benefits of the programme. Two recent works analyse randomised interventions designed to increase take-up of the Supplemental Nutrition Assistance Program (SNAP) in the US (Finkelstein and Notowidigdo 2019) and all social benefits in France (Castell et al. 2022). Both use a variation of an information-only $vs.$ human assistance compared with a control group. Both find that human assistance increase take-up. While Finkelstein and Notowidigdo (2019) find that providing targeted information also increases take-up, Castell et al. (2022) find no such effect in their setting. Finkelstein and Notowidigdo (2019) further show that the compliers are initially better off than the control group, thus leaving out those who need it most or are more likely to benefit from it.

On the importance of high quality research designs This experiment was designed and analysed with great care. I used block-randomisation to ensure balance between groups, gain precision and be able to test the presence of heterogeneity. I pre-registered the design, main outcomes and estimation methods on Social Science Registry before having access to the data. The benefit of using a randomised experiment are particularly salient in this work. Without random assignment, the next best alternative to estimate the effect of the programme on participants yields estimates of the opposite sign excluding the experimental estimates from the confidence interval. This result is worrisome considering the lack of experimental evidence on such policies, the vulnerability of the target audience and the public spending. For instance, the review of Bergemann and Van Den Berg (2008) on ALMPs for women in Europe conclude that they are very effective but only 4 out of 39 included research use a randomised experiment, two of which show negative results. The evolution of the literature on modern difference-in-differences raises additional doubts on the reliability of these estimates (J. Roth et al. 2023).

Opportunity costs of public funds In this experiment, policymakers revealed their willingness to get to lift these families out of poverty, spending roughly € 2800 per participants, setting an employment target of 10 pp for the success. They were ready to support 10 single parents so that one would get a job on average, implicitly accepting a cost of € 28 000 per expected job. This exceeds the total employer cost of one year full-time job at the minimum wage (which is approximately €22,000 per year) while relying on highly uncertain effects. And in fact, the programme did not increase employment and increased social transfers instead. This echoes one of the arguments motivating the "*Territoire Zéro Chômeur*" project, where long-term unemployed are offered permanent contracts with full time minimum wage. This high stake experiment was not randomised but the statistical service of the Ministry of Labor (DARES) used matching to evaluate the effects and find significant positive effects on participants' employment and well-being (DARES 2021). Conversely, Kasy and Lehner (2023) evaluate a similar policy in Germany and use a very high standard research design. They look at the individual average treatment effect using a randomised-match-pair design, evaluate the aggregate effect using a pre-registered synthetic control at the municipality level and comparison to individual in control municipalities. The latter allows to identify spillovers. They find positive impacts of program participation on economic and non-economic wellbeing, but not on physical health or preferences. At the municipality level, they find a large reduction of long-term unemployment, and no negative employment spillovers.

It seems important to highlight the risk taken in pursuing activation policies and to question the motivations and justifications for these choices, considering the costs involved, the risk of failure, and the little consideration for high quality design. Could different use of this same budget reach policymakers objectives ? Could one do better ? For instance, several recent studies support the idea that a significant monetary transfer produces lasting positive effects on poverty (Jones and Marinescu 2022) and children's education (Barr, Eggleston, and Smith 2022) in very different contexts. Most recently, a large scale unconditional cash transfers randomised experiment in the US showed that it increases mothers spendings for and time spend with their children (Gennetian et al. 2022). However, these amounts must be sufficient to meet the needs of families, otherwise, such interventions have no lasting effect (Jaroszewicz et al. 2022). Overall, social assistance programmes in high-income countries are insufficient in preserving the health and well-being of socio-economically disadvantaged populations, indicating that the scope and generosity of existing programmes fall short in compensating for the negative consequences associated with poverty (Shahidi et al. 2019).

'*From welfare-to-work – and worries*' While the previous presidency of Emmanuel Macron has been analysed as a continuation of the *flexicurity à la Française* started by Nicolas Sarkozy (Gazier 2019; Askenazy 2022), his second term is strongly influenced by the *"Welfare chauvinism"* pressure of Radical right populists (Rinaldi and Bekker 2021). In the European parliament, these new movements have changed the content of debates on social policies. They adopt a very different position on Welfare States from traditional Right wing parties in that they use it as mean to redefine "*the people*" - "*insiders*" and "*outsiders*", "*deserving*" and "*underserving*" - through rights and obligations.

The new welfare reform is set to make social support *mandatory* and put the Employment agency in charge of it, with large sanctions. In light of these results, this policy is not only regressive, it goes against a large and growing body of evidences[52] to which this research add yet another null effect. In particular, the analyse of the effects of sanctions for RSA recipients by Chareyron, Le Gall, and L'Horty (2022) find increased registration to Employment agency and social support but also increased non-take-up of RSA. It also dismiss alternative use of public funds that showed promising results. For instance, recent randomised experiments in Europe indicate that removing *workfare obligations* can improve labour market outcomes (Verlaat et al. 2021), or be as effective as intensive monitoring while enhancing trust in public institutions (Betkó 2023), or that job guarantee improve individual welfare, reduce long-term unemployment without negative spillovers (Kasy and Lehner 2023).

A quarter century after "Anti-poverty for families the next century: From welfare-to-work – and worries" of Ellwood (2000) in the *Journal of Economic Perspectives*, it is worrying, indeed, that question for the next 25 years turned into: *Welfare-to-what ?*

[52] Control and sanctions may expedite return to employment but at the cost of reduced wage and job quality (Arni, Lalive, and Van Ours 2013). Existing literature suggests that harsher control can direct poor individuals toward disability pensions or leave them without income (McCRATE and SMITH 1998; Dwyer et al. 2020; Morescalchi and Paruolo 2020; de Gendre, Schurer, and Zhang 2022).

Appendix

A The Reliance experiment

A.I Recruiting participants

A) Invitation letter and invitation leaflet for the first cohort

Figure A.7: Model letter for recruiting the first cohort

Laxou, le 15 Février 2018

MODELE

Dossier suivi par
ARELIA – Dispositif RELIANCE
Tél : [redacted]

Tout sur le RSA en Meurthe et Moselle :
www.insertion.meurthe-et-moselle.fr

Objet : RSA – Information collective RELIANCE

Madame,

Vous êtes bénéficiaire du Revenu de Solidarité Active (RSA).

La Caisse d'Allocations Familiales, le Conseil Départemental de Meurthe-et-Moselle ainsi que la Caisse des Dépôts et Consignation mettent en place une **action d'accompagnement destinée aux chefs de familles monoparentales**. Dans ce cadre, vous avez été identifié(e) pour y participer.

Vous êtes invité(e) à une réunion d'information collective présentant cette action intitulée RELIANCE dont l'objectif est de favoriser, à terme, votre accès à un emploi ou une formation. Au cours de cette rencontre, un temps vous sera réservé afin d'évoquer votre situation, vos projets et vos modalités d'organisation.

Celle-ci aura lieu le :

Date : **Lundi 19 Mars 2018 à 09H30 à 11H00**

Lieu : **ARELIA RELIANCE**
9-11 rue Robert Schuman, 3ème étage
54500 VANDOEUVRE LES NANCY

(Voir plan au dos)

Merci de vous organiser pour vous rendre disponible. Néanmoins, ***en cas de difficultés de garde****, nous pouvons vous accueillir en présence de vos enfants.*

Si vous ne pouvez pas venir à cette réunion, nous vous demandons de téléphoner au [redacted] **réception de cette lettre**, pour nous en informer.

Dans l'attente de vous rencontrer, je vous prie de recevoir, Madame, l'expression de nos salutations distinguées.

Pour le Président du Conseil Départemental et par délégation,

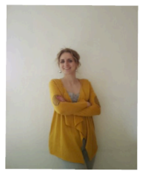
[redacted]
Responsable du Service Economie Solidaire et Insertion

Reliance est un dispositif porté par l'association Arélia, en partenariat avec Ulis et Ecoval

Figure A.8: Presentation leaflet 1/2

Figure A.9: Presentation leaflet 2/2

ATELIER : ACCES AUX DROITS/ MODE DE GARDE/ RAPPORT A SOI/ SANTE



Accès aux droits et au numérique : faire le point sur l'ouverture des droits et actualisation informatique

Recherche de mode de garde (Branche famille CAF)

Vie quotidienne et organisation : comment appréhender le changement à venir et trouver des solutions adaptées

Rapport à soi et aux autres : (parentalité, conjugalité, féminité, masculinité, santé)
comment être en harmonie avec soi et les autres au sein du changement.

ATELIER : CONSTRUCTION DU PROJET PROFESSIONNEL/ PROJET DE VIE

Détermination ou émergence d'un projet personnel (connaissance de soi, plan d'actions)

Phase d'exploration (découverte des métiers et formations)

Immersion en milieu professionnel (stage, mise à disposition, PMSMP)

Recherche d'emploi intensive et/ ou formation professionnelle

ATELIER RECS : RESEAU D'ECHANGES, DE COMPETENCES ET DE SAVOIRS/ CITOYENNETE ET BIEN-ETRE

RECS :

- Valoriser les compétences et les savoirs dans le cadre d'échange
- Favoriser l'entraide, la cohésion et le savoir-être ensemble

Atelier citoyenneté : être acteur dans la société

Atelier bien-être : prendre soin de soi (relaxation, gestion du stress, socio esthéticienne, art thérapie….)

Atelier créatif : développement des capacités, valorisation par le biais d'activités ludiques et manuelles

Version nov 2018   Ne pas jeter sur la voie publique

B) Invitation letter for the second cohort

Figure A.10: Letter in the 2019 and subsequent cohorts

Laxou, le 09 Novembre 2018

Madame XXXXXXXX XXXX
XXXXXXX
XXXXXXX

54000 NANCY

Dossier suivi par
ARELIA – Dispositif RELIANCE
Tél : [redacted]

Tout sur le RSA en Meurthe et Moselle :
www.insertion.meurthe-et-moselle.fr

Objet : RSA – Information collective RELIANCE

«Civ»,

La Caisse d'Allocations Familiales, le Conseil Départemental de Meurthe-et-Moselle ainsi que la Caisse des Dépôts et Consignation organisent une **action d'accompagnement destinée aux chefs de familles monoparentales** appelée RELIANCE. Dans ce cadre, vous avez été identifié(e) pour y participer.

L'objectif de cette action d'insertion est de vous soutenir dans la construction et la réalisation d'un projet professionnel personnalisé. Elle comprend également de nombreux ateliers organisés autour de l'accès aux droits, à la santé, à la citoyenneté, à la garde d'enfants…

Une information détaillée du dispositif vous sera présenté le :

Date : **Lundi 10 Décembre 2018 de 09H30 à 11H00**

Lieu : **ARELIA RELIANCE**
9-11 rue Robert Schuman, 3ème étage
54500 VANDOEUVRE LES NANCY

(Voir plan au dos)

Je vous informe que **votre participation à cette réunion s'inscrit dans vos obligations relatives à la législation du dispositif RSA.**

Merci de vous organiser pour vous rendre disponible. Néanmoins, **en cas de difficultés de garde**, nous pouvons vous accueillir en présence de vos enfants.

Si vous ne pouvez pas venir à cette réunion, nous vous demandons de téléphoner au [redacted] **réception de cette lettre**, pour nous en informer.

Dans l'attente de vous rencontrer, je vous prie de recevoir, Madame, l'expression de nos salutations distinguées.

Pour le Président du Conseil Départemental et par délégation,

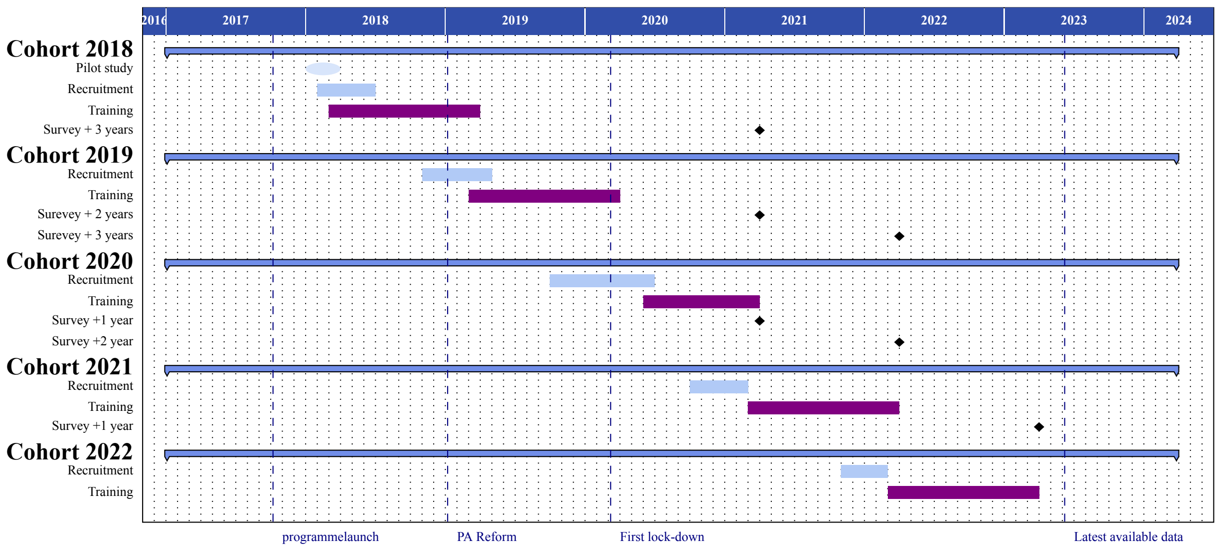
[redacted]
Responsable du Service Economie Solidaire et Insertion

Reliance est un dispositif porté par l'association Arélia, en partenariat avec Ulis et Ecoval

A.II Implementation details

A) Timeline of the experiment

Figure A.11: Timeline of the experiment



B) Adaptation of the programme during the pandemic

The Covid-19 epidemic and the sanitary restrictions implemented in France from March 2020 onwards had a significant impact on the course of support and, more importantly, on the living conditions, prospects, and opportunities for this particularly vulnerable population.

The year 2020 was marked by the impact of lockdown measures, curfews, and activity restrictions in certain sectors, to varying extents. The exceptional measures taken by authorities to support businesses and households likely mitigated the recessive effects of the health crisis to some extent. However, the economy has been durably affected, and vulnerable populations have been particularly impacted (Duvoux and Lelièvre 2021).

The first lockdown was implemented just as the participants of the 2020 cohort were starting their support journey. One challenge for the Reliance team was to maintain a connection despite the lockdown. To achieve this, a coordinator was assigned to regularly call all participants and answer their inquiries during extended time slots, including evenings and weekends. Social media, particularly a Facebook group, was utilised, with the coordinator regularly sharing diverse content ranging from recipes and activities to job-related quizzes, documentation, and information, etc.

The Reliance team also identified limited access to digital resources for some participants. Consequently, computer equipment (laptops, tablets) was provided for free to certain participants with the support of the Caf and the Departmental Council. Social workers also arranged for food parcels to be delivered to homes and printed important documents for beneficiaries who no longer had access to printing facilities. A mask manufacturing plant offered employment opportunities for participants to make masks, and volunteer sewing workshops for mask-making were organised.

The qualitative evaluation highlighted that contact was generally well maintained, and those interviewed felt supported. However, activities initiated before the lockdown were put on hold (such as driver's licence exams, workplace internships, access to rights, etc.), and the Reliance team noticed a greater level of "disengagement," particularly less participation in collective activities despite the implemented safety measures. FORS (2020) notes explanations from participants, such as fear of rejoining groups in the current health context or overwhelming anxiety. Overall, the lockdown period was anxiety-inducing for Reliance's supported individuals, who once again found themselves isolated, potentially exacerbating their difficulties. However, FORS also highlights that this period helped re-motivate some participants despite the context.

B Data and presentation of the variables used in the analysis

B.I Data sources

Data quality and attrition I observed all households that received at least one CAF payment in the year, but some data are collected more frequently than others and some come from other administrations (*e.g.* taxes or unemployment benefits). Those who apply for RSA, PA, or the disability allowance for adults (AAH) must report their incomes and situation every quarters. Most variables come from these quarterly reports after statistical consolidations[53]. Families in our sample have been registered for RSA for at least two years and are, therefore, used to fill these quarterly declarations. On average we have complete data for 92% the period after randomisation. Only 4.1% files are lost. Figure B.12 compares the attrition rate across cohorts and samples and shows that the missing income or missing files are very similar for the encouraged and control groups, whereas those excluded by social workers exhibit higher attrition patterns. In figure B.13, I estimate the effect of encouragement on the probability of observing incomes each month and show the event study-plot[54]. Overall, there are no differential attrition between the treatment arms although there is a small bump of higher reporting at the early stage of the programme.

Sample restrictions The block sizes vary a lot and some of them have four observations or less. There are few blocks with less than 2 observations in each treatment arms. In order to compute cluster-robust standard errors at the block level, I drop observations from these blocks from the analysis. This deletion is innocuous and reduce the full database from 103066 individual $\times$ month observations to 102749. Only 1 treated household is dropped.

In addition, I restrict the window of analysis to the timeframe with overlapping relative times since randomisation to avoid composition effects. I can only observe the 2018 cohort from -13 month before randomisation. For the 2021 cohort, data go up to 30 months after randomisation. Estimations over 4 cohorts are always over this interval. I also look at the effects over the first three cohorts over a longer period.

Finally, the 2022 cohort includes 127 households from the control groups of the previous cohorts. For instance, in the 2018 cohort, 24 were re-sampled, among which 12 were invited with the fourth cohort, 5 of which enrolled. In the main analysis, the observations of the encouragement group among those re-sampled are dropped, creating a slight compositional change in the analysis. There data are analysed with the 2022 cohort. Those resampled but remaining in the control group are duplicated in the dataset, once in their original cohort and once in the control group of the 2022 cohort.

Non-response: casewise deletion These administrative data have little problem of missing variables apart from attrition which is independent from the treatment. There are however some that have incoherent values and missing observations coded 99999 in original files when, for instance, parents stop for a quarter to send their income. So far, I simply delete missing cases.

[53] For each month, there are three files corresponding to three extraction delays: FR1, FR2, and FR6, corresponding to 1, 2, and 6 months of recall. The quality of information about beneficiaries and their rights to benefits increases with extraction delay and these are certified data The later the files, the better they include "latecomers". "Latecomers" occur because parents did not send all information on time or because there were administrative controls, delays in mail treatment, errors and other procedures.

[54] This model is estimated exactly like other models for intention to treat analysis. See section IV for more details.

B.II Description of the main variables of interest

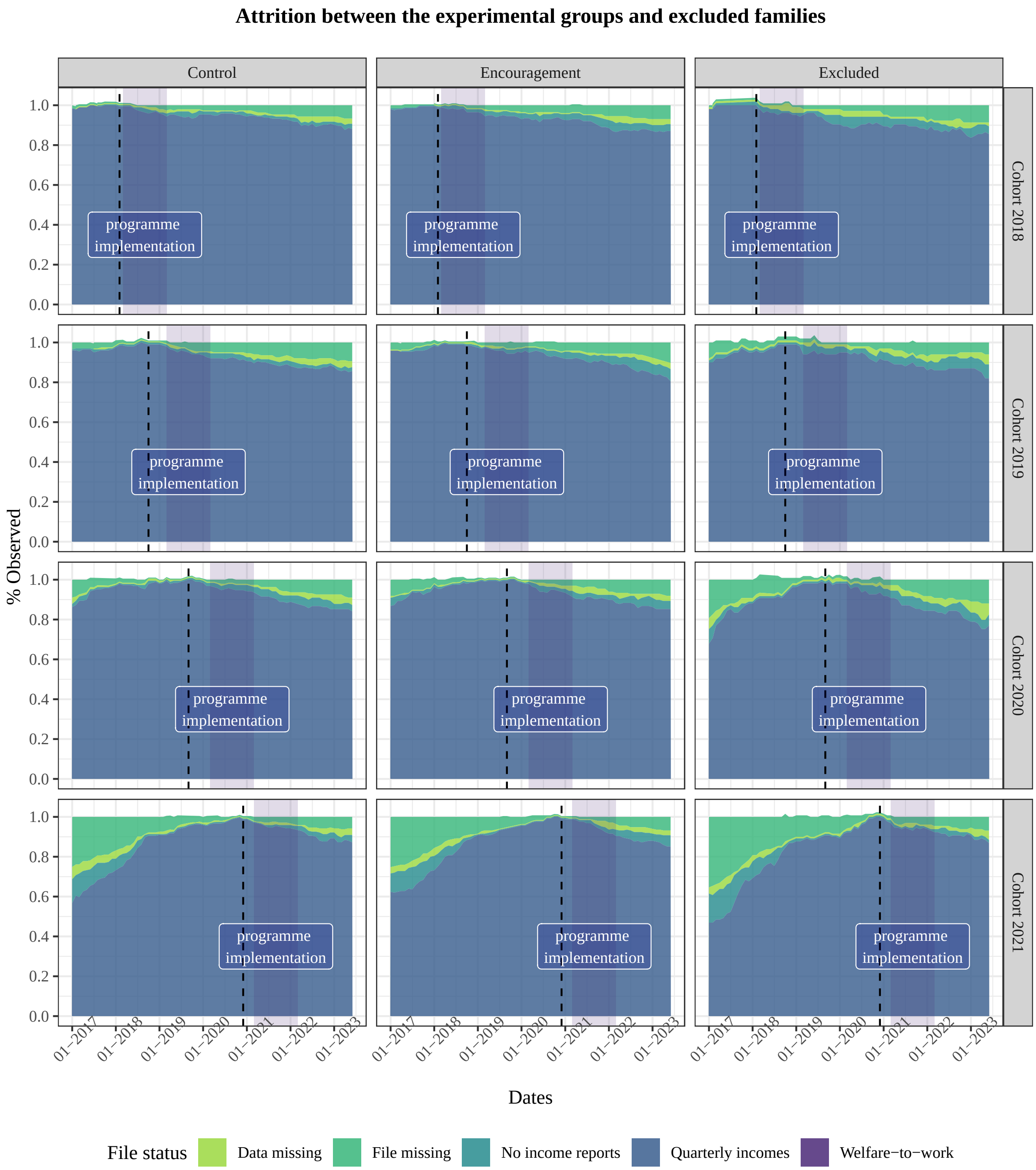
Table B.4: Definition of main outcome variables

Outcomes	Variables	Variable names	Description
Main outcomes	Labour Income >0	income_pos	*Dummy for positive labour income of the sampled parent. Built from quarterly reports for RSA and/or PA (MTACMMEX|MTACMONX)*
	Receive Rsa	out_RSA	*Dummy for receiving RSA (at least € 15 paid)*
	Monthly Total Household's Incomes	incomeRUC	*Total disposable incomes of the household, including social transfers and spouse incomes. Computed by the national family allowance fund and actualised using the INSEE CPI for bottom quintile of the income distribution.*
	Monthly Household's Incomes Per Cu	RUC	*Total household incomes weighted by family size and structure. First adult has weight 1, any other adult or child over 14 has weight .5, additional children under 14 have .3 weights. Single parents have an additional .2 weight. Computed by the national family allowance fund and actualised using the INSEE CPI for bottom quintile of the income distribution.*
	Monthly Total Social Transfers	MTPRESVE	*Total transfers from the Family allowance fund. Computed by the national family allowance fund actualised using the INSEE CPI for bottom quintile of the income distribution.*
Alternative measurements	'In Work'	Employment	*Dummy that equals one if parents 'main' occupation is 'in work' according to INSEE 7 categories employment classification. Coded from ACTRESPD, imputed by social workers when files are updated. Low reliability on the timing*
	Incomes Above Rsa Threshold	out_RSAtight	*Dummy for exiting RSA because earnings are above the eligibility threshold. Computed from motives for RSA ineligibility or payment cancelling.*
	Individual Earnings	INCOME	*Individual earnings of the sampled parent computed from quarterly reports*

B.III Attrition

A) Descriptive statistics

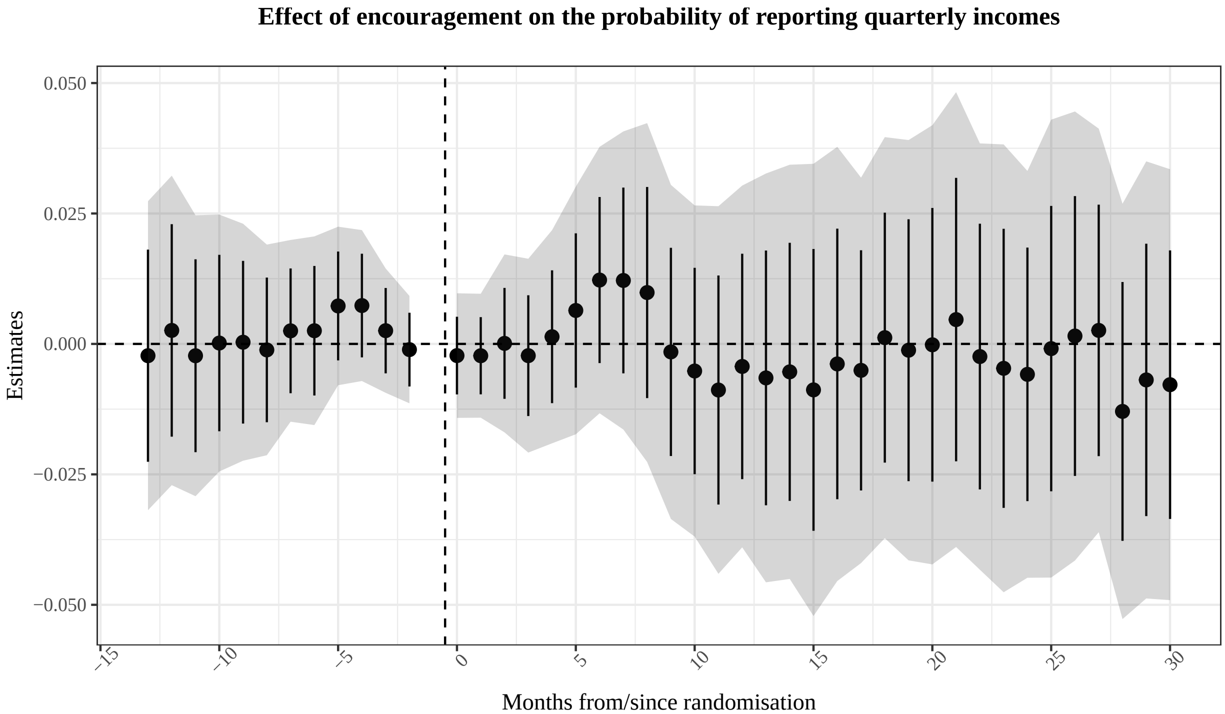
Figure B.12: Share of income data available across time and cohorts



Sources: ALLSTAT 2017−01−01 to 2023−06−01
Proportions of the baseline population observed or not at each date
in the experimental group and excluded families.

B) Estimating the intention-to-treat effect on the probability of reporting quarterly incomes

Figure B.13: Event study of the effect of encouragement on availability of incomes' data



Sources: ALLSTAT 2017−01−01 to 2023−06−01
Notes: The dependent variable equals 1 if we observe quartely incomes the corresponding month, 0 otherwise.
Event study with block x cohort x relative month fixed effect and centered encouragement by block−propensity scores.
Error bars, indicates 95 % CI using cluster−robust standard errors at the cohort level.
Shades indicates 95%CI adjusting for the FWER.

B.IV Other variables used in the analysis

Table B.5: Definition and label for variables used as covariates

Variables	Variable names	Description
Age	AGE	*Parent's age at the month of randomisation. Computed from Birth date in the ALLSTAT data.*
Age youngest child	youngest	*Age of the youngest child computed from ALLSTAT data.*
Age oldest child	oldest	*Age of the oldest child computed from ALLSTAT data.*
Gender	Sexe	*Gender as defined by social security numbers, original renamed from ALLSTAT data.*
N Children under 2	NBEN0A2C	*Number of children under 2 year-old in december of the year, computed by CAF.*
N Children 3 to 5	NBEN3A5C	*Number of children 3 to 5 years-old in december of the year, computed by CAF.*
STI assessment	SESIDummy	*Dummy for social workers assessment of the family's appropriatness for the program. 1 for favourable assessment, 0 for reserved.*
French citizenship	French	*Dummy for French cityzenship from ALLSTAT*
Family allowance	AF	*Dummy for receiving 'Allocations familiales', constructed from AFVERS*
Family supplement	CF	*Dummy for receiving 'Complément familiale', constructed from CFVERS*
Housing benefit	APL	*Dummy for receiving one of the three housing benefit, constructed from DROAIDLO*
Family support allowance	ASF	*Dummy for receiving 'Allocation de soutien familiale', constructed from ASFVERS*
Child support	TPA	*Dummy for receiving 'pension alimentaire',original renamed from ALLSTAT*
Education	School3	*Level of education grouped in 3 values (High/Low/Unknown): High := for High school degree , Higher education and Vocational degree ; Low:= Middle school diploma , No education , Other training or Some Middle school.*
N consumption units	NBUC	*Number of consumption units computed by Cnaf. 1 unit of for the first adult, 0.5 per additional adult or child aged 14 and over, 0.3 per child under 14 and 0.2 for a single-parent family.*

Sources: ALLSTAT. Data preparation files are available upon request.

Table B.5: Definition and label for variables used as covariates

Variables	Variable names	Description
Early childhood allowance	PAJE	*Dummy for receiving 'Préstation d'accueil du jeune enfant', constructed from PAJEVERS*

Sources: ALLSTAT. Data preparation files are available upon request.

B.V Balance check at the time of random assignment

Table B.6: Balance of main variables of interest the month before randomisation

	Control (N=828)		Encouragement (N=843)		Diff. in Means	Std. Error
	Mean	*Std. Dev.*	*Mean*	*Std. Dev.*		
Share labour income >0	0.08	0.28	0.08	0.27	-0.01	0.02
Share receive RSA	0.98	0.15	0.98	0.14	0.00	0.01
Mean monthly total household's incomes	1394.69	493.21	1398.98	497.86	4.37	24.08
Mean monthly household's incomes per CU	708.59	150.11	712.19	157.06	3.70	11.51
Mean monthly total social transfers	1297.67	497.91	1292.00	481.42	-7.10	17.99
Mean yearly taxable income N-2	1424.11	2871.59	1599.39	3315.34	178.27	206.54
Favourable assessment	0.66	0.47	0.69	0.46	0.03	0.03
Mean distance (km) to the programme	3.32	1.86	3.49	2.01	0.17	0.12
Share French	0.81	0.39	0.84	0.37	0.03	0.02
Share higher education	0.51	0.50	0.53	0.50	0.02	0.03
Share lower education	0.25	0.43	0.24	0.43	0.00	0.03
Share unknown education	0.24	0.43	0.23	0.42	-0.01	0.02
Mean age	36.04	7.95	36.14	7.75	0.08	0.46
Mean age youngest child	7.13	5.56	7.16	5.49	0.03	0.34
Mean age oldest child	11.41	6.28	11.32	6.12	-0.11	0.34
Share with children under 2	0.31	0.46	0.29	0.45	-0.02	0.03
Share with children 3 to 5	0.33	0.47	0.32	0.47	-0.01	0.03
Share with one child over 16	0.32	0.47	0.30	0.46	-0.03	0.03
Share receive family allowance	0.57	0.50	0.56	0.50	-0.01	0.01
Share receive family supplement	0.17	0.38	0.17	0.38	0.00	0.03
Share receive housing benefit	0.89	0.31	0.88	0.32	-0.01	0.02

* = p<.1, ** = p<.05, *** = p<.01
Sources: ALLSTAT, cohorts 2018 to 2021 one month before randomisation.
Notes : mean and mean differences are weighted within-block averages.
Standard errors account for block randomisation.

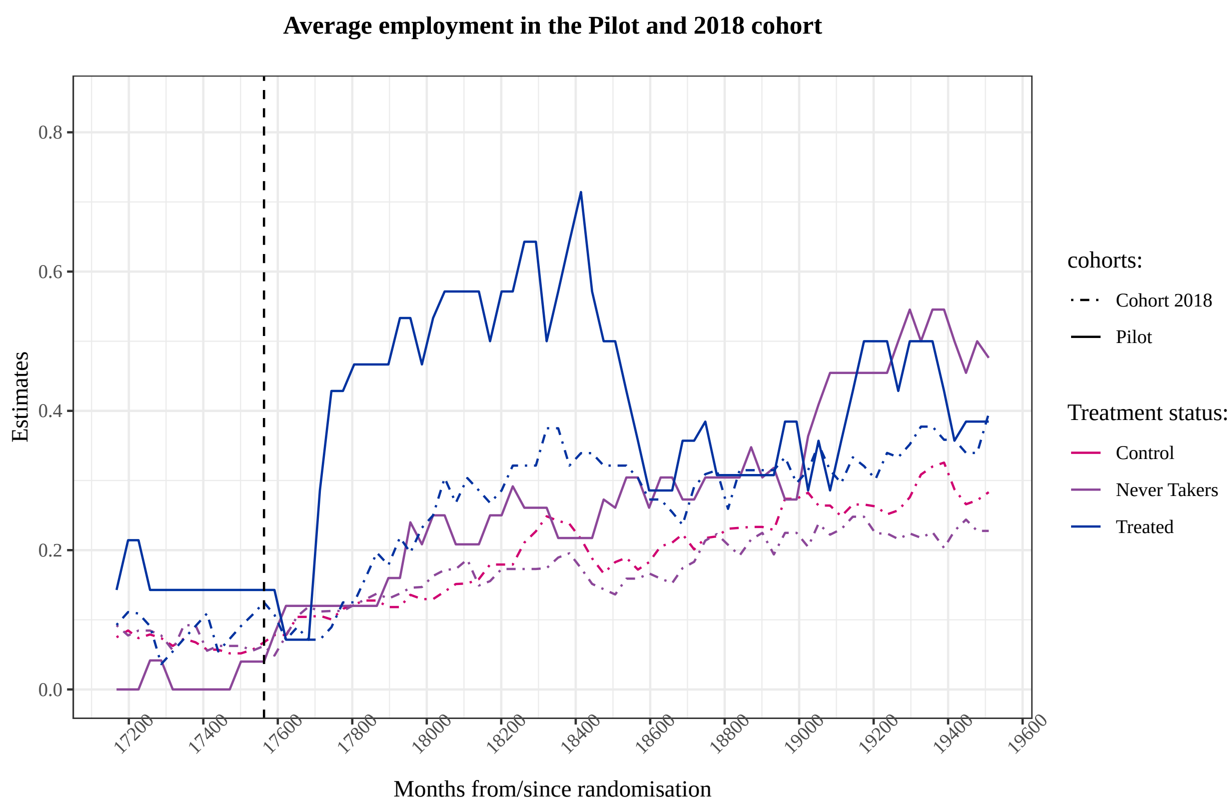
Table B.6: Balance of main variables of interest the month before randomisation

	Control (N=828)		Encouragement (N=843)		Diff. in Means	Std. Error
	Mean	*Std. Dev.*	*Mean*	*Std. Dev.*		
Share receive family support allowance	0.65	0.48	0.64	0.48	-0.01	0.02
Share Receive child support	0.21	0.40	0.21	0.40	0.00	0.02
Share receive Early childhood allowance	0.31	0.46	0.29	0.45	-0.02	0.03

* = p<.1, ** = p<.05, *** = p<.01
Sources: ALLSTAT, cohorts 2018 to 2021 one month before randomisation.
Notes : mean and mean differences are weighted within-block averages.
Standard errors account for block randomisation.

B.VI Comparison of the outcomes of the 2018 cohort and the pilot group

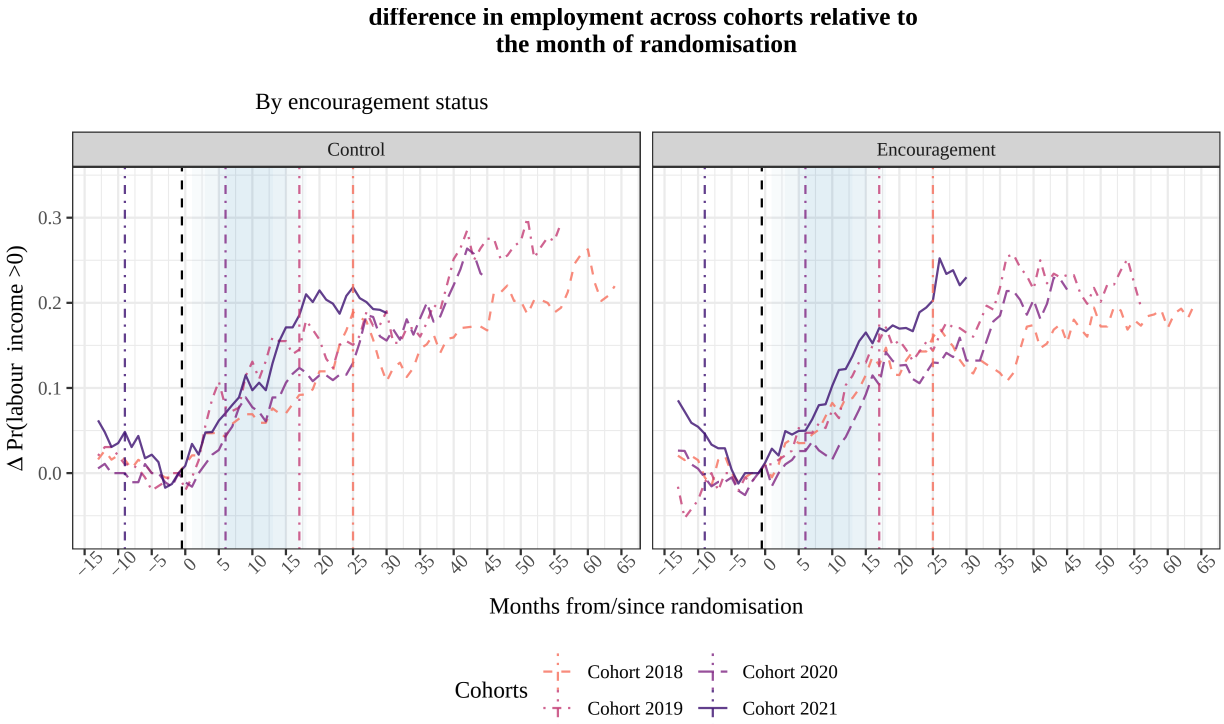
Figure B.14: Evolution of the share of parents with positive labour income in the pilot and first cohort



Sources: ALLSTAT 2017−01−01 to 2023−06−01
Notes: The dependent variable equals 1 when the person has positive labour income and 0 otherwise.

B.VII Comparison of average changes in employment in relative time across cohorts and treatment arms

Figure B.15: Never more than 25 pp increase in employment in all 4 cohorts



Sources: ALLSTAT 2017−01−01 to 2023−06−01
The dependent variable is the long difference between employment at month m and employment at the time of randomisation.
Lines indicates sample means color codded by cohort.
The dash−dot lines indicate the month of the first lock−down for each cohort.

B.VIII Comparison of average disposable income per consumption units across cohorts and treatment arms

Figure B.16: Average disposable income per consumption unit over the period

Sources: ALLSTAT 2017−01−01 to 2023−06−01 cohorts 2018 to 2021.
Points indicate the sample mean by date and randomisation group. Error bars indicates pointwise 95 % CI using .975 quantile of a normal distribution, sample size and variance.
Smooth lines are estimated using splines with 1/6 of the number of dates as degree of freedom.

C Additional estimates of the effect of the programme

C.I Aggregated treatment effect on employment

Table C.7: Aggregated estimates on labour market maret participation

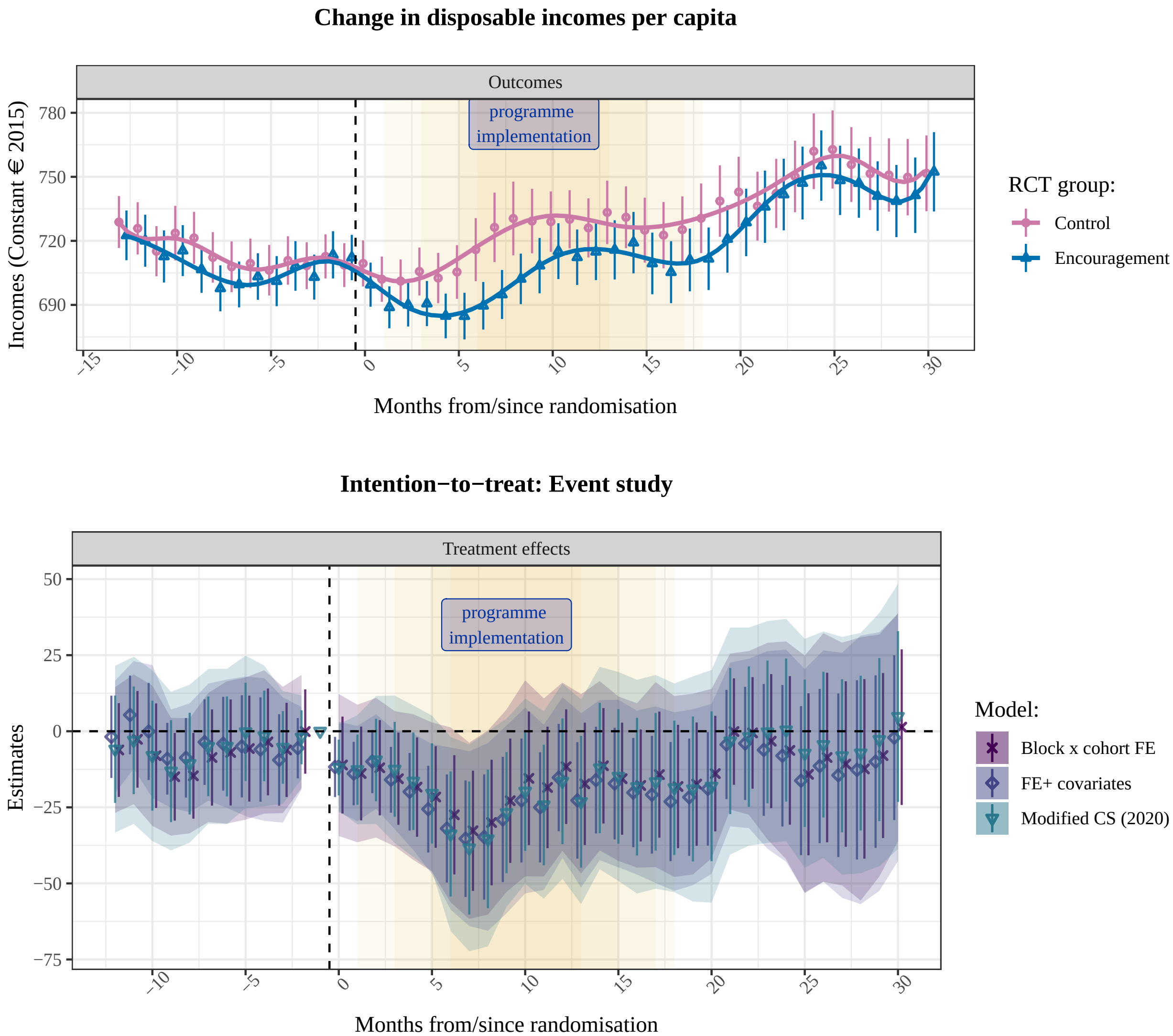
	Mean control	**OLS**		**TSLS**	
		No covariates	*Covariates*	*No covariates*	*Covariates*
[-7 ; -1 [	0.085*** (0.012)	-0.009 (0.013)	-0.006 (0.009)	-0.024 (0.034)	-0.017 (0.023)
	[0.06, 0.11]	[-0.04, 0.02]	[-0.03, 0.02]	[-0.11, 0.06]	[-0.08, 0.04]
	adj.p.val. = 0.000	*adj.p.val. = 0.927*	*adj.p.val. = 0.945*	*adj.p.val. = 0.927*	*adj.p.val. = 0.944*
[0 ; 6 [	0.113*** (0.012)	-0.014 (0.012)	-0.015 (0.010)	-0.036 (0.030)	-0.039 (0.026)
	[0.08, 0.14]	[-0.04, 0.02]	[-0.04, 0.01]	[-0.11, 0.04]	[-0.11, 0.03]
	adj.p.val. = 0.000	*adj.p.val. = 0.626*	*adj.p.val. = 0.477*	*adj.p.val. = 0.627*	*adj.p.val. = 0.473*
[6 ; 12 [	0.164*** (0.016)	-0.029** (0.014)	-0.036** (0.014)	-0.073** (0.034)	-0.093*** (0.035)
	[0.13, 0.20]	[-0.06, 0.01]	[-0.07, 0.00]	[-0.16, 0.01]	[-0.18, 0.00]
	adj.p.val. = 0.000	*adj.p.val. = 0.140*	*adj.p.val. = 0.051*	*adj.p.val. = 0.128*	*adj.p.val. = 0.039*
[12 ; 18 [	0.203*** (0.020)	-0.006 (0.016)	-0.013 (0.016)	-0.016 (0.042)	-0.032 (0.042)
	[0.16, 0.25]	[-0.05, 0.04]	[-0.05, 0.03]	[-0.12, 0.09]	[-0.14, 0.07]
	adj.p.val. = 0.000	*adj.p.val. = 0.997*	*adj.p.val. = 0.927*	*adj.p.val. = 0.997*	*adj.p.val. = 0.927*
[18 ; 24 [	0.235*** (0.021)	-0.009 (0.019)	-0.020 (0.017)	-0.022 (0.047)	-0.052 (0.042)
	[0.18, 0.29]	[-0.06, 0.04]	[-0.06, 0.02]	[-0.14, 0.10]	[-0.16, 0.06]
	adj.p.val. = 0.000	*adj.p.val. = 0.989*	*adj.p.val. = 0.675*	*adj.p.val. = 0.989*	*adj.p.val. = 0.665*
[24 ; 30 [	0.260*** (0.020)	-0.008 (0.020)	-0.014 (0.019)	-0.021 (0.049)	-0.036 (0.048)
	[0.21, 0.31]	[-0.06, 0.04]	[-0.06, 0.03]	[-0.15, 0.10]	[-0.16, 0.09]
	adj.p.val. = 0.000	adj.p.val. = 0.994	adj.p.val. = 0.940	adj.p.val. = 0.994	adj.p.val. = 0.940
Num.Obs.	28125	56749	56749	56749	56749
R2	0.027	0.108	0.280	0.106	0.274
R2 Adj.	0.027	0.067	0.245	0.064	0.239
Covariates			X		X

* $p < 0.1$, ** $p < 0.05$,*** $p < 0.01$ using point-wise p-value. Adjusted p-value and confidence intervals account for simultaneous inference using the Holm–Bonferroni correction. Standard errors are cluster-heteroskedasticity robust adjusted at the block x cohort level.

Notes: Control group means estimated using OLS with period dummies and no constant. OLS columns indicates average ITTs, TSLS columns indicate average ATTs. All models include block x cohort x relative time fixed effects and use inverse instrument propensity score weighting for double-robustness. Encouragement variable is centred by the instrument propensity score. I report the average of the F-stats for the first stages of all treatment periods. Covariates are measured at the month before random assignment and include baseline level, initial assessment, education, French citizenship, quartiles of age, receiving child support or family support allowance, housing benefits, children under 3, children 3 to 5, children older than 16, taxable income higher than median and quartile of baseline income per capita. I also include dummies for being resampled in the 2022 cohort and being encouraged. All covariates are interacted with relative time dummies to have specific effects for each period.

C.II Dynamic intention to treat on income per consumption unit

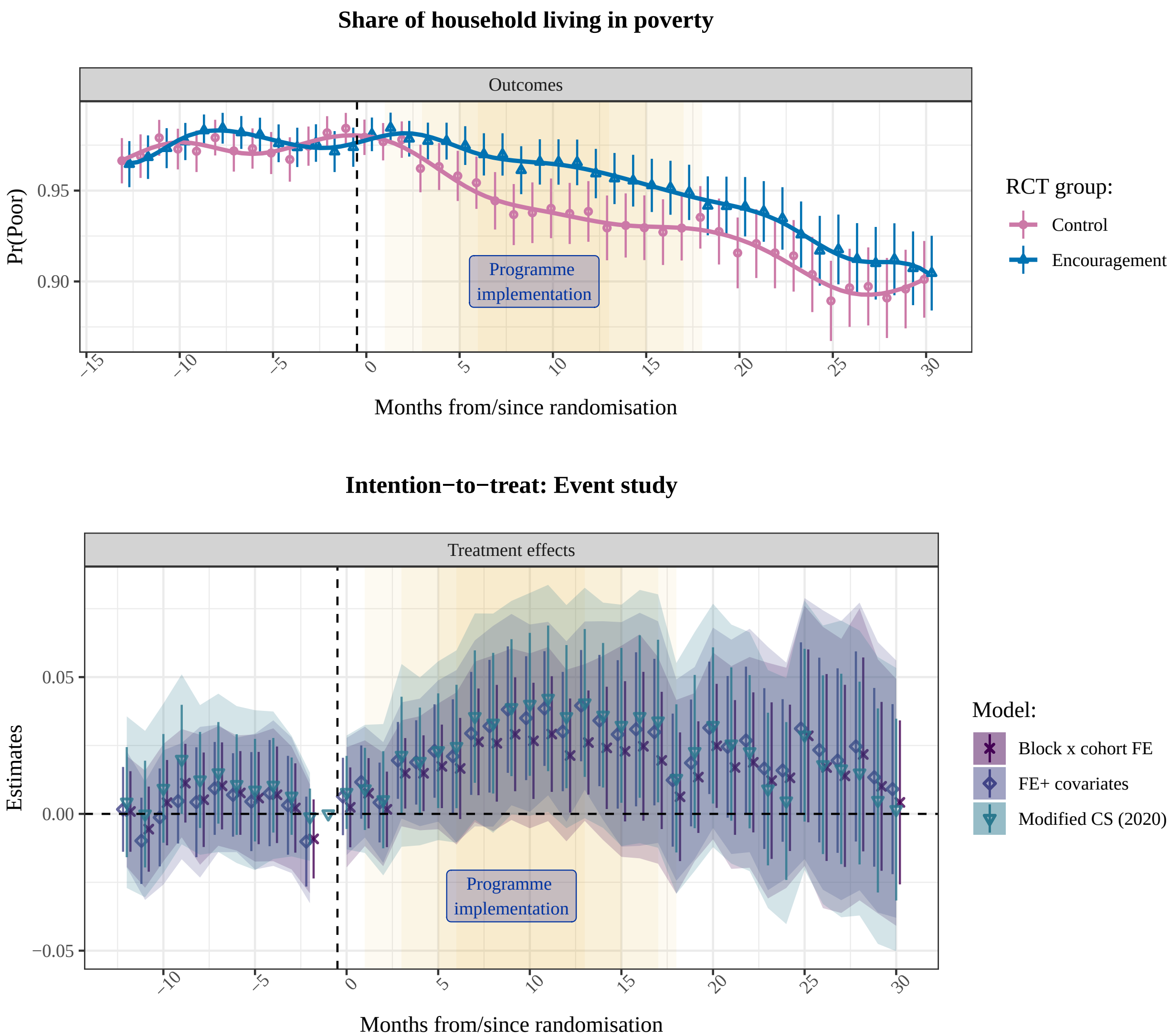
Figure C.17: Dynamic Effects of the programme on incomes per capita



Sources: ALLSTAT 2017−01−01 to 2023−06−01
Notes: The dependent variable is the disposable income per consumption unit.
Top panel:
− Points indicate simple means by encouragement status over cohorts 2018 to 2021 in relative time since randomisation.
− Error bars indicate pointwise 95% CI using simple standard errors.
− Lines are conditional means estimated with spline regressions.
Bottom panel:
− Event study with block x cohort fixed effects with/without covariates using OLS or modified Callaway Sant' Anna (2020).
− Error bars, indicates 95 % CI using cluster−robust standard errors at the block x cohort level.
− Shades indicates 95%CI adjusting for the FWER using Holm−Bonferroni correction for OLS and wild cluster bootstrap for modified Callaway and Sant' Anna (2020).
Covariates are measured at the month before random assignment and include baseline level, initial assessment, education, French citizenship, quartiles of age, receiving child support or family support allowance, housing benefits, children under 3, children 3 to 5, children older than 16, taxable income higher than median and quartile of baseline income per capita.
All covariates are interacted with relative time dummies to have specific effects for each period.
For OLS models, covariates also include dummies for being resampled in the 2022 cohort and being encouraged, interacted with relative time dummies. For the modified CS model, observations of the late encouraged group are dropped.

C.III Dynamic intention to treat on the risk of poverty

Figure C.18: Dynamic Effects of the programme on the probability of living in poverty

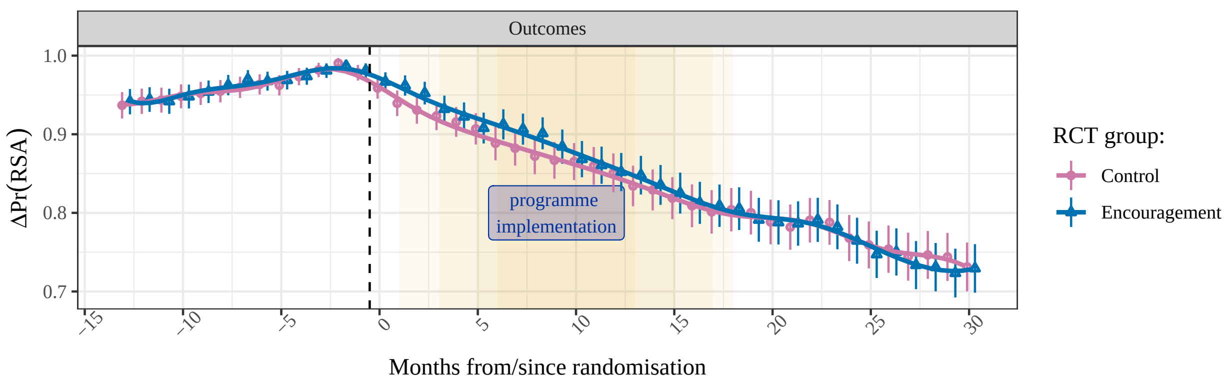


Sources: ALLSTAT 2017−01−01 to 2023−06−01
Notes: The dependent variable is a dummy for disposable income per capita lower than the poverty line.
Top panel:
− Points indicate simple means by encouragement status over cohorts 2018 to 2021 in relative time since randomisation.
− Error bars indicate pointwise 95% CI using simple standard errors.
− Lines are conditional means estimated with spline regressions.
Bottom panel:
− Event study with block x cohort fixed effects with/without covariates using OLS or modified Callaway Sant' Anna (2020).
− Error bars, indicates 95 % CI using cluster−robust standard errors at the block x cohort level .
− Shades indicates 95%CI adjusting for the FWER with Holm−Bonferroni correction for OLS and wild cluster bootstrap for modified Callaway and Sant' Anna (2020).
Covariates are measured at the month before random assignment and include initial assessment, education, French citizenship, quartiles of age, receiving child support or family support allowance, housing benefits, children under 3, children 3 to 5, children older than 16, taxable income higher than median and quartile of baseline income per capita.
All covariates are interacted with relative time dummies to have specific effects for each period.
For OLS models, covariates also include dummies for being resampled in the 2022 cohort and being encouraged, interacted with relative time dummies. For the modified CS model, observations of the late encouraged group are dropped.

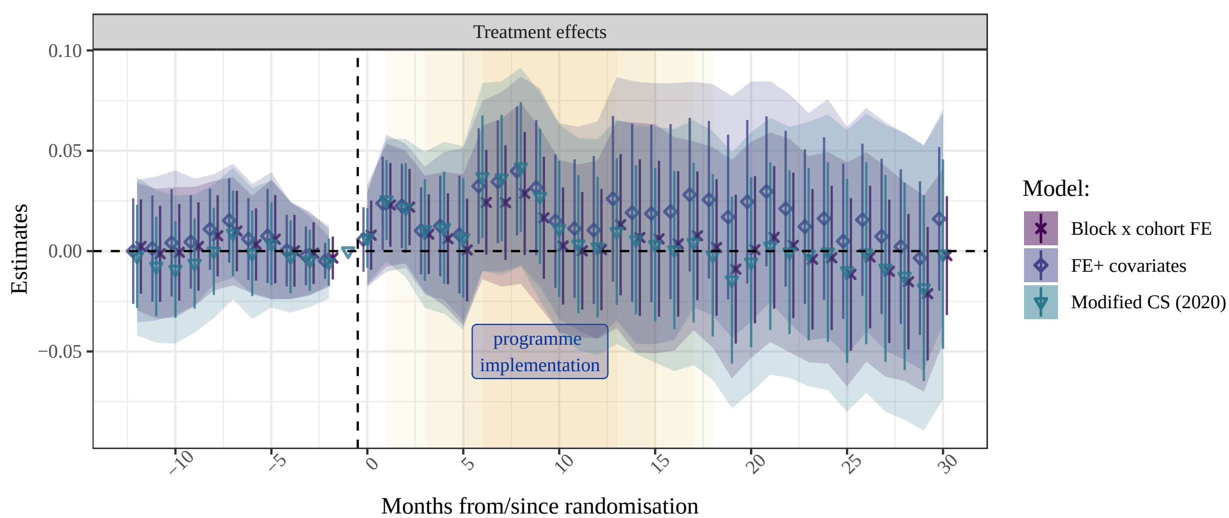
C.IV Effect of the programme on RSA and PPA take-up

Figure C.19: Intention-to-treat effects of the programme on the probability of receiving RSA payments

Change in the share receiving RSA

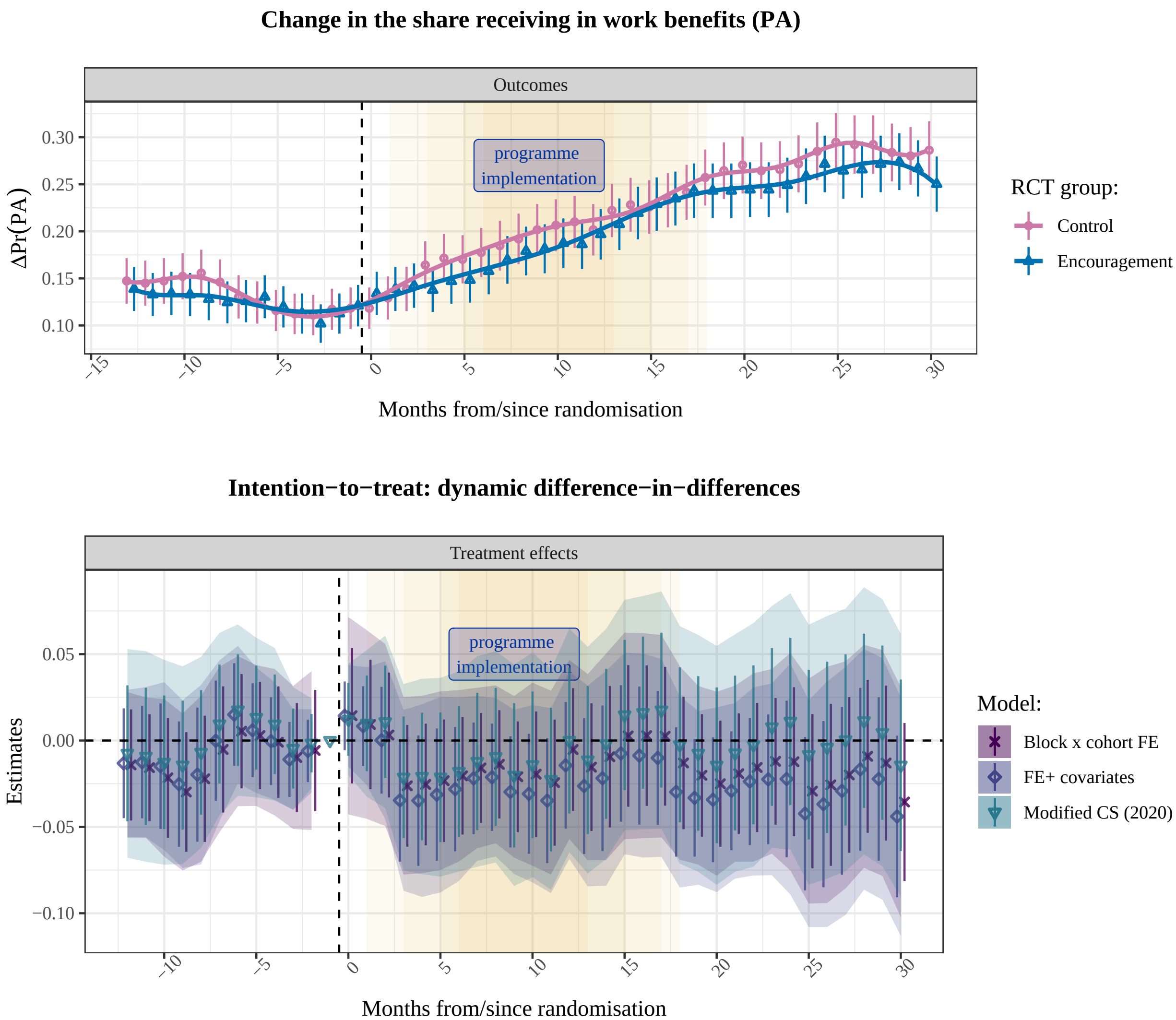


Intention−to−treat: dynamic difference−in−differences



Sources: ALLSTAT 2017−01−01 to 2023−06−01
Notes: The dependent variable is a dummy for receiving RSA.
Top panel:
− Points indicate simple means by encouragement status over cohorts 2018 to 2021 in relative time since randomisation.
− Error bars indicate pointwise 95% CI using simple standard errors.
− Lines are conditional means estimated with spline regressions.
Bottom panel:
− Event study with block x cohort fixed effects with/without covariates using OLS, demeaned encouragement and inverse propensity score weighting or modified Callaway Sant' Anna (2020) estimator (See section IV).
− Error bars, indicates 95 % CI using cluster−robust standard errors at the block x cohort level.
− Shades indicates 95%CI adjusting for the FWER using Holm−Bonferroni correction for OLS and wild cluster bootstrap for modified Callaway and Sant' Anna (2020).
Covariates are measured at the month before random assignment and include baseline level, initial assessment, education, French citizenship, quartiles of age, receiving child support or family support allowance, housing benefits, children under 3, children 3 to 5, children older than 16, taxable income higher than median and quartile of baseline income per capita.
All covariates are interacted with relative time dummies to have specific effects for each period.
For OLS models, covariates also include dummies for being resampled in the 2022 cohort and being encouraged, interacted with relative time dummies. For the modified CS model, observations of the late encouraged group are dropped.

Figure C.20: Intention-to-treat effects of the programme on the probability of receiving in work benefits (PA)

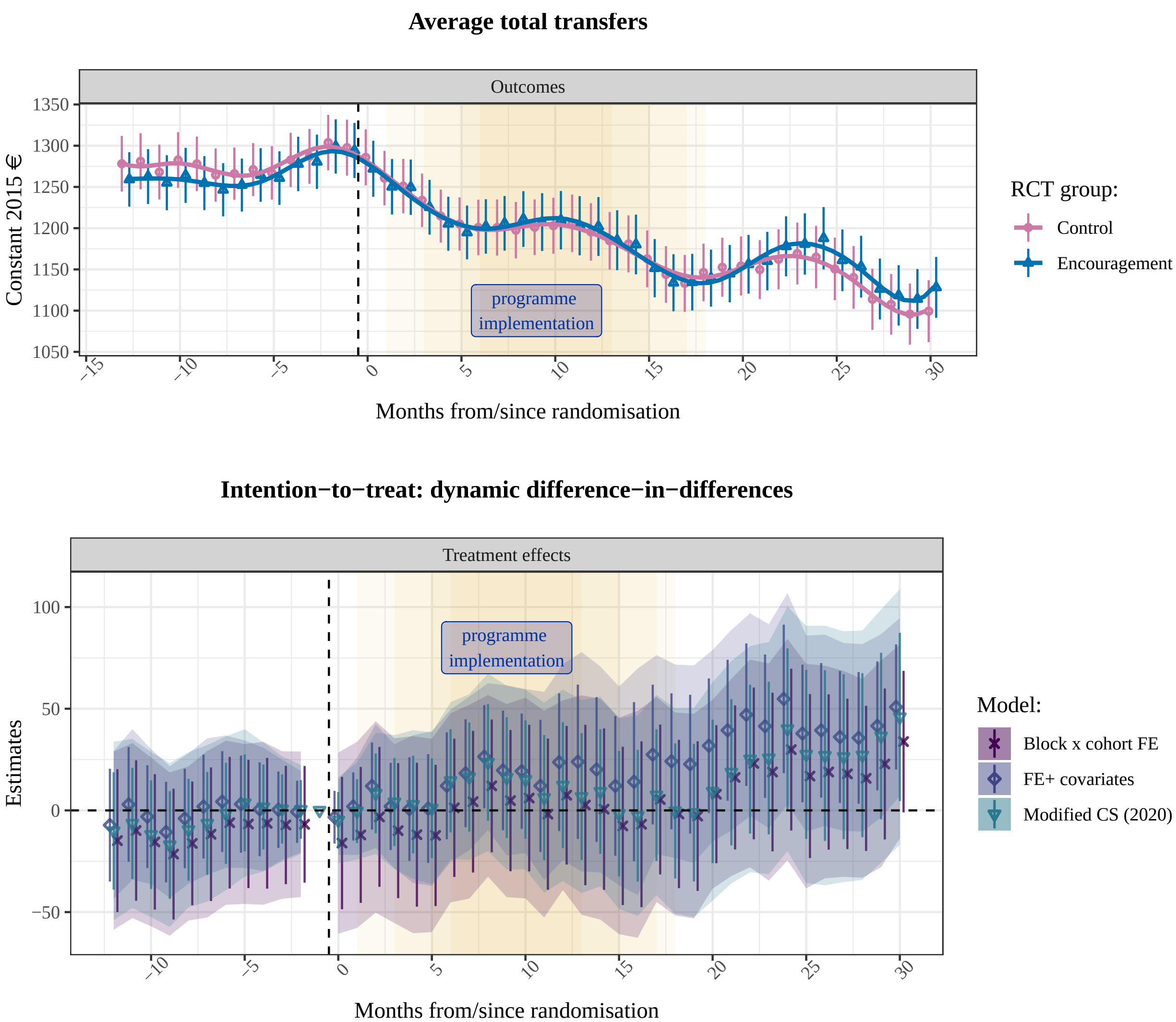


Sources: ALLSTAT 2017−01−01 to 2023−06−01
Notes: The dependent variable is a dummy for receiving PA, the in−work benefit.
Top panel:
− Points indicate simple means by encouragement status over cohorts 2018 to 2021 in relative time since randomisation.
− Error bars indicate pointwise 95% CI using simple standard errors.
− Lines are conditional means estimated with spline regressions.
Bottom panel:
− Event study with block x cohort fixed effects with/without covariates using OLS, demeaned encouragement and inverse propensity score weighting or modified Callaway Sant' Anna (2020) estimator (See section IV).
− Error bars, indicates 95 % CI using cluster−robust standard errors at the block x cohort level.
− Shades indicates 95%CI adjusting for the FWER using Holm−Bonferroni correction for OLS and wild cluster bootstrap for modified Callaway and Sant' Anna (2020).
Covariates are measured at the month before random assignment and include baseline level, initial assessment, education, French citizenship, quartiles of age, receiving child support or family support allowance, housing benefits, children under 3, children 3 to 5, children older than 16, taxable income higher than median and quartile of baseline income per capita.
All covariates are interacted with relative time dummies to have specific effects for each period.
For OLS models, covariates also include dummies for being resampled in the 2022 cohort and being encouraged, interacted with relative time dummies. For the modified CS model, observations of the late encouraged group are dropped.

C.V Effects of the programme on total transfers from the Family allowance fund

A) Dynamic intention to treat on total amount of cash transfers

Figure C.21: Intention-to-treat effects of the programme on total amount of transfers from the Family allowance fund



Sources: ALLSTAT 2017−01−01 to 2023−06−01
Notes: The dependent variable is the total amount of cash transfers, in constant 2015 symbol("\xa0") .
Top panel:
− Points indicate simple means by encouragement status over cohorts 2018 to 2021 in relative time since randomisation.
− Error bars indicate pointwise 95% CI using simple standard errors.
− Lines are conditional means estimated with spline regressions.
Bottom panel:
− Event study with block x cohort fixed effects with/without covariates using OLS, demeaned encouragement and inverse propensity score weighting or modified Callaway Sant' Anna (2020) estimator (See section IV).
− Error bars, indicates 95 % CI using cluster−robust standard errors at the block x cohort level.
− Shades indicates 95%CI adjusting for the FWER using Holm−Bonferroni correction for OLS and wild cluster bootstrap for modified Callaway and Sant' Anna (2020).
Covariates are measured at the month before random assignment and include baseline level, initial assessment, education, French citizenship, quartiles of age, receiving child support or family support allowance, housing benefits, children under 3, children 3 to 5, children older than 16, taxable income higher than median and quartile of baseline income per capita.
All covariates are interacted with relative time dummies to have specific effects for each period.
For OLS models, covariates also include dummies for being resampled in the 2022 cohort and being encouraged, interacted with relative time dummies. For the modified CS model, observations of the late encouraged group are dropped.

B) Aggregated effect on total social transfers

Table C.8: Aggregated effects of the programme on total transfers from the family allowance fund

	Mean control	OLS		TSLS	
		No covariates	*Covariates*	*No covariates*	*Covariates*
[-7 ; -1 [	1280.2*** (52.1)	-7.5 (14.5)	1.5 (9.5)	-19.3 (37.7)	4.0 (24.7)
	[1168.5, 1391.9]	[-43.1, 28.2]	[-22.4, 25.5]	[-112.0, 73.4]	[-58.2, 66.3]
	adj.p.val. = 0.000	*adj.p.val. = 0.964*	*adj.p.val. = 1.000*	*adj.p.val. = 0.965*	*adj.p.val. = 1.000*
[0 ; 6 [	1242.3*** (51.3)	-10.9 (16.1)	2.4 (9.0)	-28.0 (41.8)	6.3 (23.3)
	[1132.6, 1352.1]	[-50.5, 28.6]	[-20.3, 25.2]	[-130.8, 74.7]	[-52.3, 64.9]
	adj.p.val. = 0.000	*adj.p.val. = 0.898*	*adj.p.val. = 0.999*	*adj.p.val. = 0.902*	*adj.p.val. = 0.999*
[6 ; 12 [	1202.4*** (50.0)	4.5 (16.6)	18.0 (12.8)	11.4 (42.4)	46.2 (32.7)
	[1095.3, 1309.5]	[-36.2, 45.2]	[-14.1, 50.0]	[-93.0, 115.8]	[-36.0, 128.4]
	adj.p.val. = 0.000	*adj.p.val. = 0.998*	*adj.p.val. = 0.476*	*adj.p.val. = 0.998*	*adj.p.val. = 0.474*
[12 ; 18 [	1167.4*** (50.4)	0.3 (18.3)	20.2 (16.8)	0.7 (47.0)	51.9 (43.0)
	[1059.4, 1275.4]	[-44.7, 45.3]	[-22.1, 62.5]	[-114.8, 116.3]	[-56.2, 160.0]
	adj.p.val. = 0.000	*adj.p.val. = 1.000*	*adj.p.val. = 0.623*	*adj.p.val. = 1.000*	*adj.p.val. = 0.620*
[18 ; 24 [	1155.9*** (49.5)	10.2 (17.5)	34.5** (16.1)	26.2 (44.5)	88.2** (41.0)
	[1049.8, 1261.9]	[-32.6, 53.1]	[-5.9, 74.8]	[-83.3, 135.6]	[-15.1, 191.4]
	adj.p.val. = 0.000	*adj.p.val. = 0.939*	*adj.p.val. = 0.123*	*adj.p.val. = 0.939*	*adj.p.val. = 0.123*
[24 ; 30 [	1128.9*** (48.9)	20.4 (17.8)	41.0*** (15.2)	51.9 (45.5)	104.4** (39.6)
	[1024.1, 1233.7]	[-23.3, 64.0]	[2.8, 79.2]	[-60.1, 163.9]	[4.7, 204.1]
	adj.p.val. = 0.000	adj.p.val. = 0.586	adj.p.val. = 0.031	adj.p.val. = 0.593	adj.p.val. = 0.036
Num.Obs.	28963	58486	58486	58486	58486
R2	0.011	0.525	0.721	0.524	0.718
R2 Adj.	0.010	0.503	0.708	0.502	0.705
Covariates			X		X

* p < 0.1, ** p < 0.05,*** p < 0.01 using point-wise p-value. Adjusted p-value and confidence intervals account for simultaneous inference using the Holm–Bonferroni correction. Standard errors are cluster-heteroskedasticity robust adjusted at the block x cohort level.

Notes: Control group means estimated using OLS with period dummies and no constant. OLS columns indicates average ITTs, TSLS columns indicate average ATTs. All models include block x cohort x relative time fixed effects and use inverse instrument propensity score weighting for double-robustness. Encouragement variable is centred by the instrument propensity score. I report the average of the F-stats for the first stages of all treatment periods.

Covariates are measured at the month before random assignment and include baseline level, initial assessment, education, French citizenship, quartiles of age, receiving child support or family support allowance, housing benefits, children under 3, children 3 to 5, children older than 16, taxable income higher than median and quartile of baseline income per capita. I also include dummies for being resampled in the 2022 cohort and being encouraged. All covariates are interacted with relative time dummies to have specific effects for each period.

C) Aggregated effect on the probability of living in poverty

Table C.9: Aggregated effects of the programme on poverty

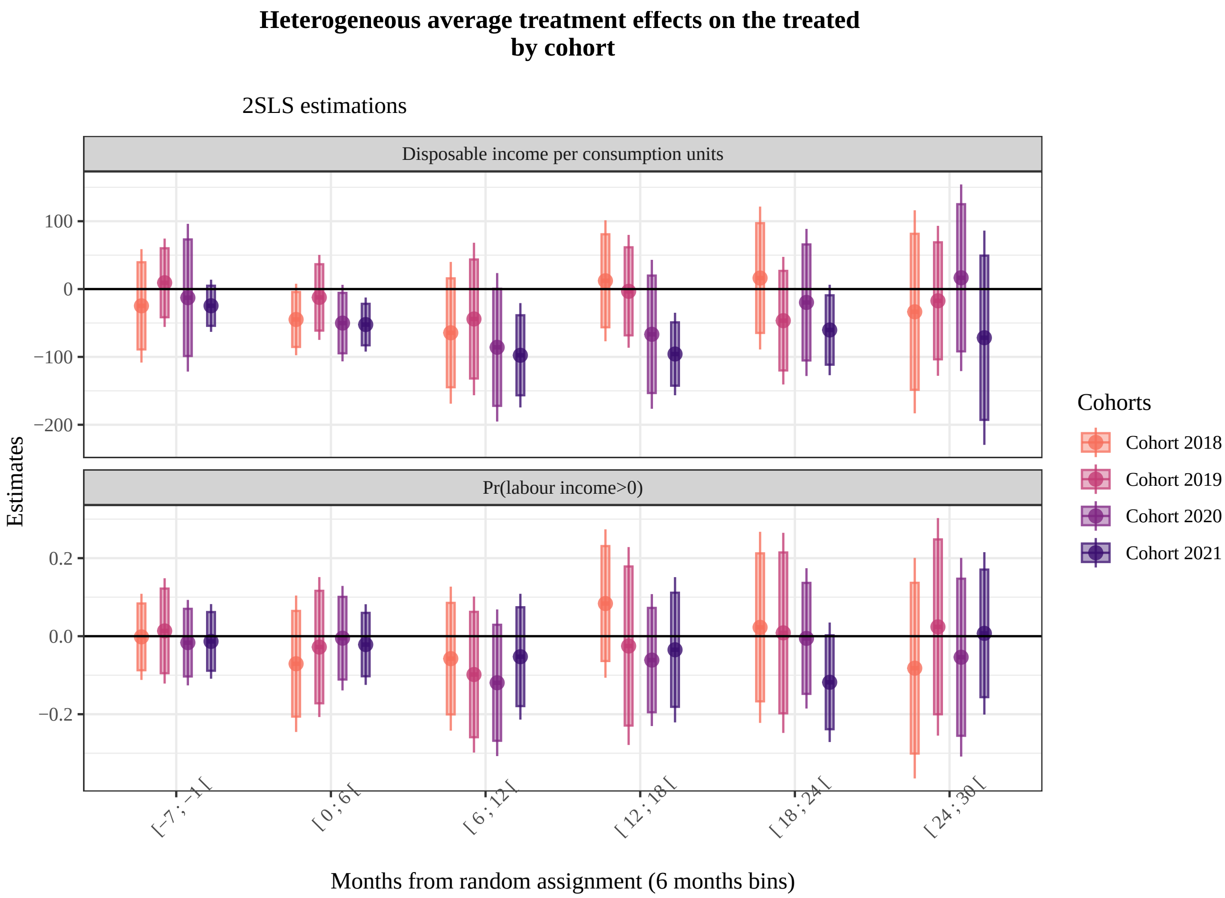
	Mean control	OLS		TSLS	
		No covariates	*Covariates*	*No covariates*	*Covariates*
[-7 ; -1 [	0.97*** (0.01)	0.00 (0.01)	0.00 (0.01)	0.01 (0.02)	0.01 (0.02)
	[0.96, 0.99]	[-0.01, 0.02]	[-0.01, 0.02]	[-0.03, 0.05]	[-0.04, 0.06]
	adj.p.val. = 0.000	*adj.p.val. = 0.980*	*adj.p.val. = 0.988*	*adj.p.val. = 0.981*	*adj.p.val. = 0.992*
[0 ; 6 [	0.97*** (0.00)	0.01* (0.01)	0.01* (0.01)	0.03* (0.01)	0.03** (0.02)
	[0.96, 0.98]	[0.00, 0.02]	[0.00, 0.03]	[-0.01, 0.06]	[-0.01, 0.07]
	adj.p.val. = 0.000	*adj.p.val. = 0.296*	*adj.p.val. = 0.211*	*adj.p.val. = 0.276*	*adj.p.val. = 0.206*
[6 ; 12 [	0.94*** (0.01)	0.03*** (0.01)	0.03*** (0.01)	0.07*** (0.02)	0.08*** (0.02)
	[0.92, 0.96]	[0.00, 0.05]	[0.01, 0.05]	[0.01, 0.12]	[0.02, 0.14]
	adj.p.val. = 0.000	*adj.p.val. = 0.016*	*adj.p.val. = 0.007*	*adj.p.val. = 0.013*	*adj.p.val. = 0.006*
[12 ; 18 [	0.93*** (0.01)	0.02** (0.01)	0.03*** (0.01)	0.06** (0.02)	0.08*** (0.03)
	[0.91, 0.95]	[0.00, 0.05]	[0.00, 0.06]	[0.00, 0.12]	[0.01, 0.15]
	adj.p.val. = 0.000	*adj.p.val. = 0.091*	*adj.p.val. = 0.016*	*adj.p.val. = 0.078*	*adj.p.val. = 0.013*
[18 ; 24 [	0.92*** (0.01)	0.02 (0.01)	0.02* (0.01)	0.04 (0.03)	0.05* (0.03)
	[0.90, 0.95]	[-0.01, 0.04]	[-0.01, 0.05]	[-0.03, 0.10]	[-0.02, 0.12]
	adj.p.val. = 0.000	*adj.p.val. = 0.477*	*adj.p.val. = 0.254*	*adj.p.val. = 0.459*	*adj.p.val. = 0.245*
[24 ; 30 [	0.90*** (0.01)	0.02 (0.01)	0.02 (0.01)	0.04 (0.03)	0.05 (0.03)
	[0.86, 0.93]	[-0.02, 0.05]	[-0.01, 0.05]	[-0.04, 0.13]	[-0.03, 0.13]
	adj.p.val. = 0.000	adj.p.val. = 0.666	adj.p.val. = 0.469	adj.p.val. = 0.657	adj.p.val. = 0.475
Num.Obs.	28700	57927	57927	57927	57927
R2	0.013	0.067	0.095	0.065	0.088
R2 Adj.	0.012	0.025	0.051	0.022	0.044
Covariates			X		X

* p < 0.1, ** p < 0.05,*** p < 0.01 using point-wise p-value. Adjusted p-value and confidence intervals account for simultaneous inference using the Holm–Bonferroni correction. Standard errors are cluster-heteroskedasticity robust adjusted at the block x cohort level.
Notes: Control group means estimated using OLS with period dummies and no constant. OLS columns indicates average ITTs, TSLS columns indicate average ATTs. All models include block x cohort x relative time fixed effects and use inverse instrument propensity score weighting for double-robustness. Encouragement variable is centred by the instrument propensity score. I report the average of the F-stats for the first stages of all treatment periods.
Covariates are measured at the month before random assignment and include baseline level, initial assessment, education, French citizenship, quartiles of age, receiving child support or family support allowance, housing benefits, children under 3, children 3 to 5, children older than 16, taxable income higher than median and quartile of baseline income per capita. I also include dummies for being resampled in the 2022 cohort and being encouraged. All covariates are interacted with relative time dummies to have specific effects for each period.

D Estimations of heterogeneous treatment effects

D.I Treatment effect heterogeneity on employment by cohort

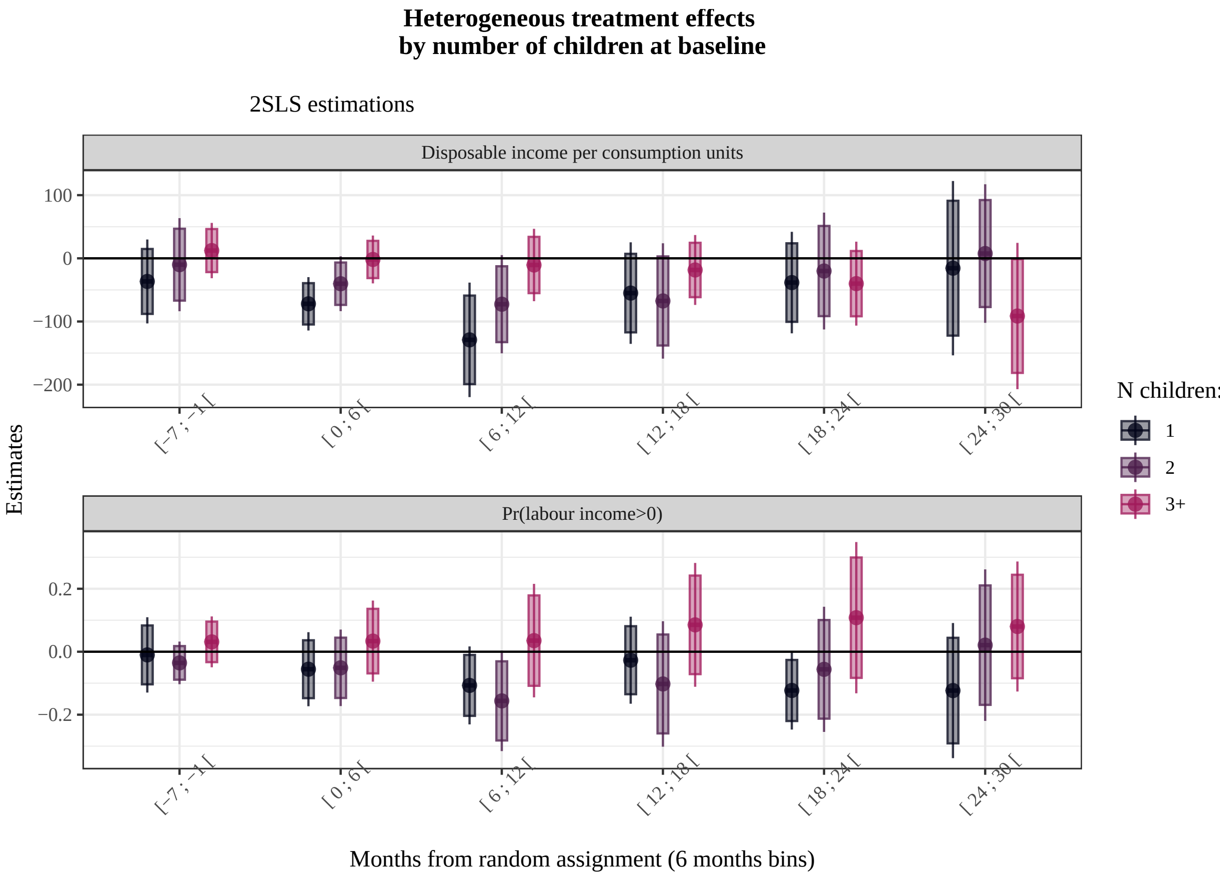
Figure D.22: Heterogeneity on effect on labour market participation across cohorts



Sources: ALLSTAT 2017−01−01 to 2023−06−01 cohorts 2018 to 2021.
Notes: Estimations using TSLS with cohort x participation x period dummies
instrumented by the demeaned instrument x cohort x period dummies.
All models include blocks x cohort x relative months fixed effects. Covariates are measured at the month
before random assignment and include baseline level, initial assessment, education, French citizenship,
quartiles of age, receiving child support or family support allowance, housing benefits, children under 3,
children 3 to 5, children older than 16, taxable income higher than median and quartile of baseline income per capita.
All covariates are interacted with relative time dummies to have specific effects for each period.
Standard errors are cluster−heteroskedasticity robusts adjusted at the block x cohort level.
− Left panel: The dependent variable is the monthly disposable income per capita.
− right panel: The dependent variable is a dummy for positive labour incomes.
− Error bars indicate 95 % pointwise confidence intervals and extended lines account for FWER by subgroup.

D.II Treatment effect heterogeneity by number of children at baseline

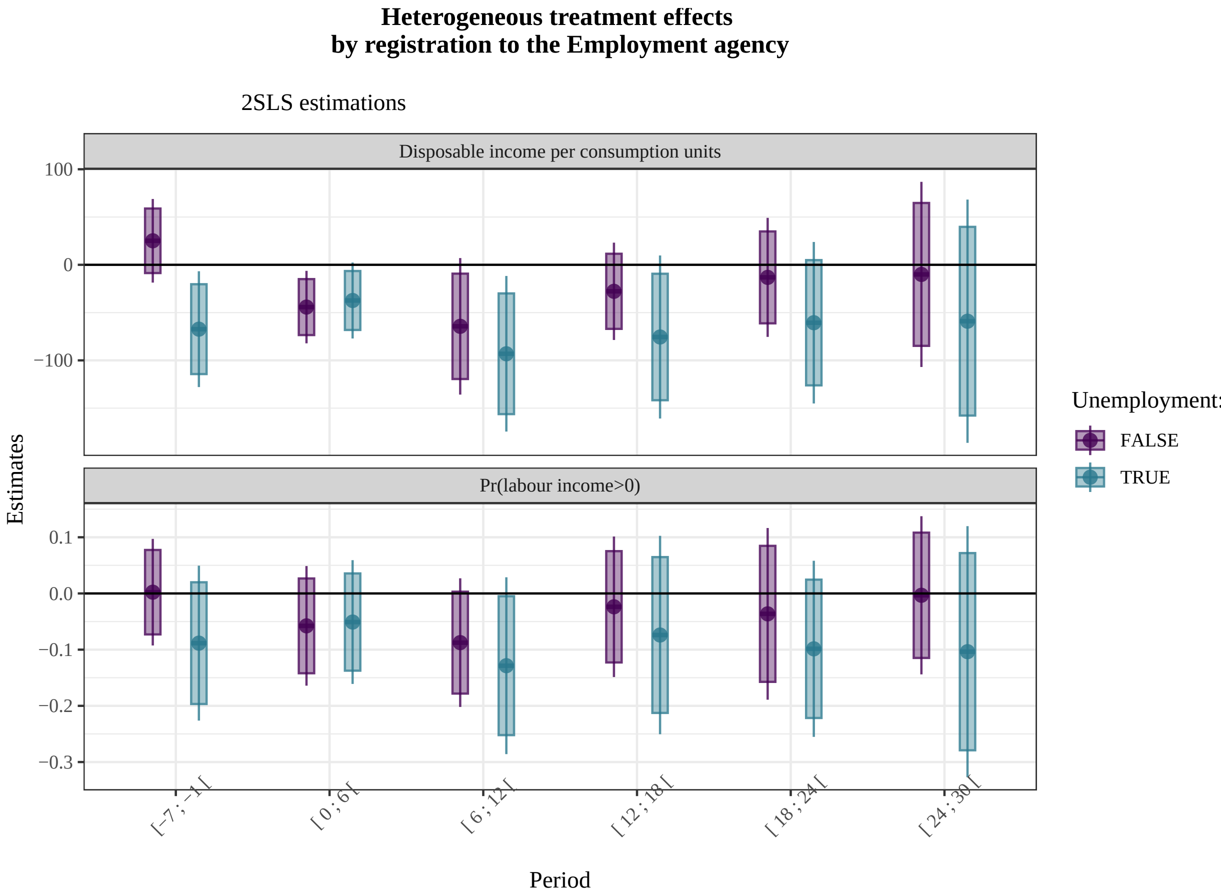
Figure D.23: Heterogenous treatment effects on disposable incomes per capita by number of children



Sources: ALLSTAT 2017−01−01 to 2023−06−01 cohorts 2018 to 2021.
Notes: Estimations using TSLS with children x participation x period dummies
instrumented by the demeaned instrument x children x period dummies.
All models include blocks x cohort x relative months fixed effects. Covariates are measured at the month
before random assignment and include baseline level, initial assessment, education, French citizenship,
quartiles of age, receiving child support or family support allowance, housing benefits, children under 3,
children 3 to 5, children older than 16, taxable income higher than median and quartile of baseline income per capita.
All covariates are interacted with relative time dummies to have specific effects for each period.
Standard errors are cluster−heteroskedasticity robusts adjusted at the block x cohort level.
− Left panel: The dependent variable is the monthly disposable income per capita.
− right panel: The dependent variable is a dummy for positive labour incomes.
− Error bars indicate 95 % pointwise confidence intervals and extended lines account for FWER by subgroup.

D.III Treatment effect heterogeneity on disposable incomes by registration to the Employment agency

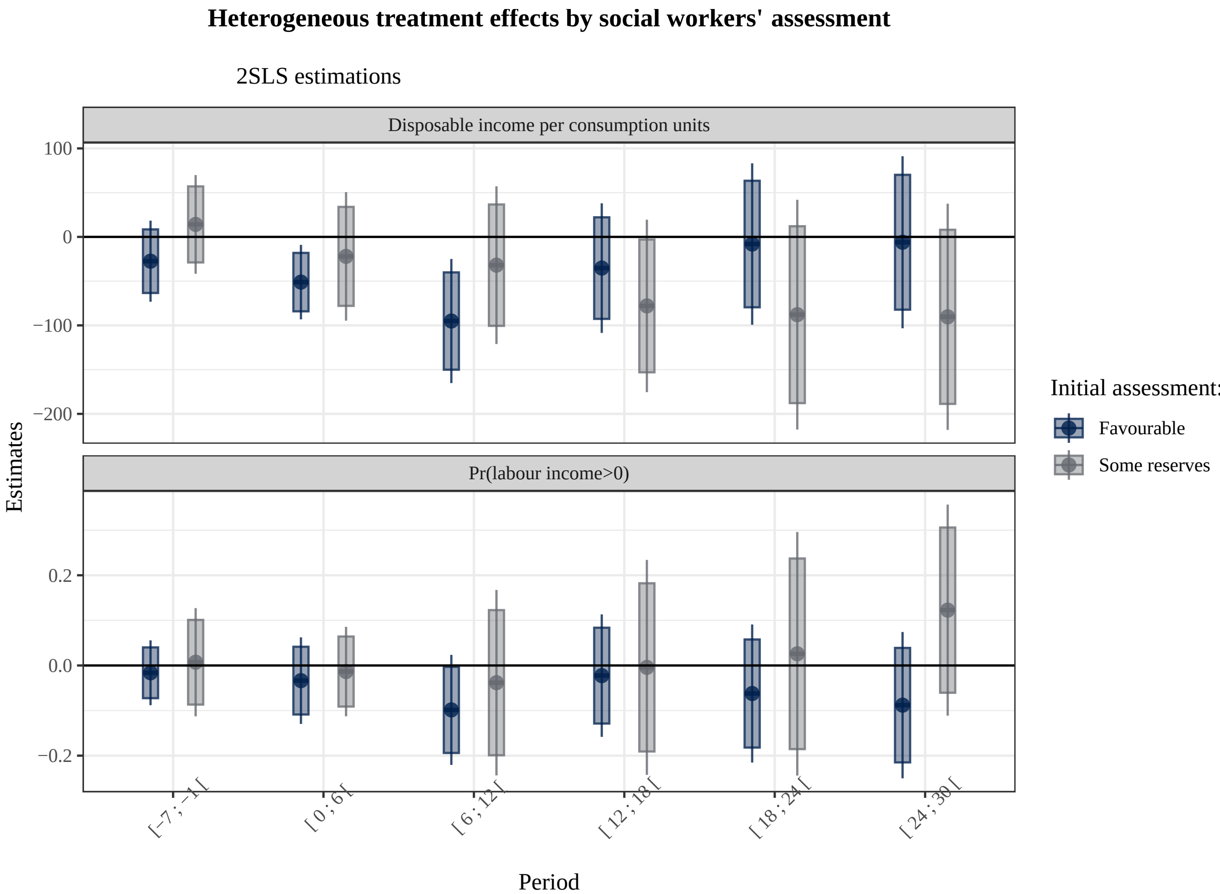
Figure D.24: Heterogenous treatment effects on employment by unemployment status



Sources: ALLSTAT 2017−01−01 to 2023−06−01 cohorts 2018 to 2021.
Notes: Estimations using TSLS with unemployment x participation x period dummies instrumented by the demeaned instrument x unemployment x period dummies.
All models include blocks x cohort x relative months fixed effects. Covariates are measured at the month before random assignment and include baseline level, initial assessment, education, French citizenship, quartiles of age, receiving child support or family support allowance, housing benefits, children under 3, children 3 to 5, children older than 16, taxable income higher than median and quartile of baseline income per capita.
All covariates are interacted with relative time dummies to have specific effects for each period.
Standard errors are cluster−heteroskedasticity robusts adjusted at the block x cohort level.
− Left panel: The dependent variable is the monthly disposable income per capita.
− right panel: The dependent variable is a dummy for positive labour incomes.
− Error bars indicate 95 % pointwise confidence intervals and extended lines account for FWER by subgroup.

D.IV Treatment effect heterogeneity by social workers’ initial assessment

Figure D.25: Heterogenous treatment effects by social workers’ initial assessment



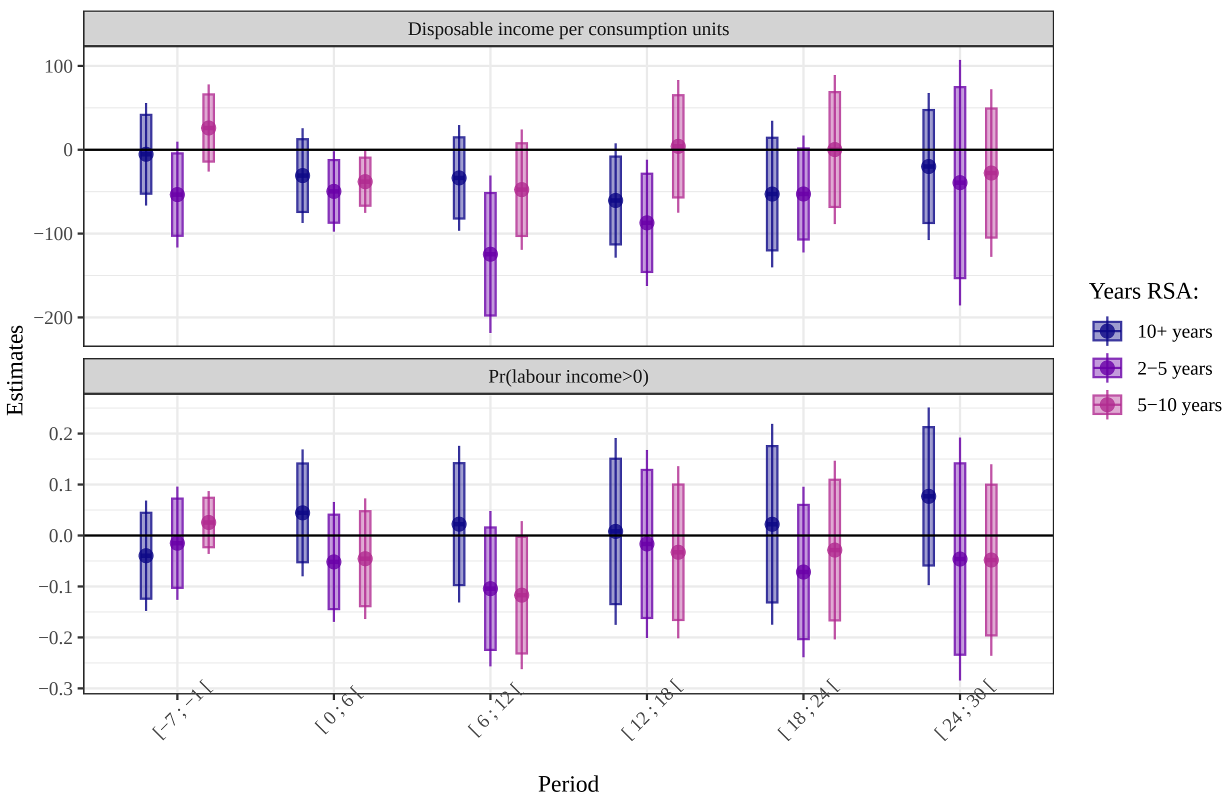
Sources: ALLSTAT 2017−01−01 to 2023−06−01 cohorts 2018 to 2021.
Notes: Estimations using TSLS with assessment x participation x period dummies instrumented by the demeaned instrument x assessment x period dummies.
All models include blocks x cohort x relative months fixed effects. Covariates are measured at the month before random assignment and include baseline level, initial assessment, education, French citizenship, quartiles of age, receiving child support or family support allowance, housing benefits, children under 3, children 3 to 5, children older than 16, taxable income higher than median and quartile of baseline income per capita.
All covariates are interacted with relative time dummies to have specific effects for each period.
Standard errors are cluster−heteroskedasticity robusts adjusted at the block x cohort level.
− Left panel: The dependent variable is the monthly disposable income per capita.
− right panel: The dependent variable is a dummy for positive labour incomes.
− Error bars indicate 95 % pointwise confidence intervals and extended lines account for FWER by subgroup.

D.V Treatment effect heterogeneity by number of years receiving RSA

Figure D.26: Heterogenous treatment effects by social workers' initial assessment

Heterogeneous treatment effects by number of years receiving RSA

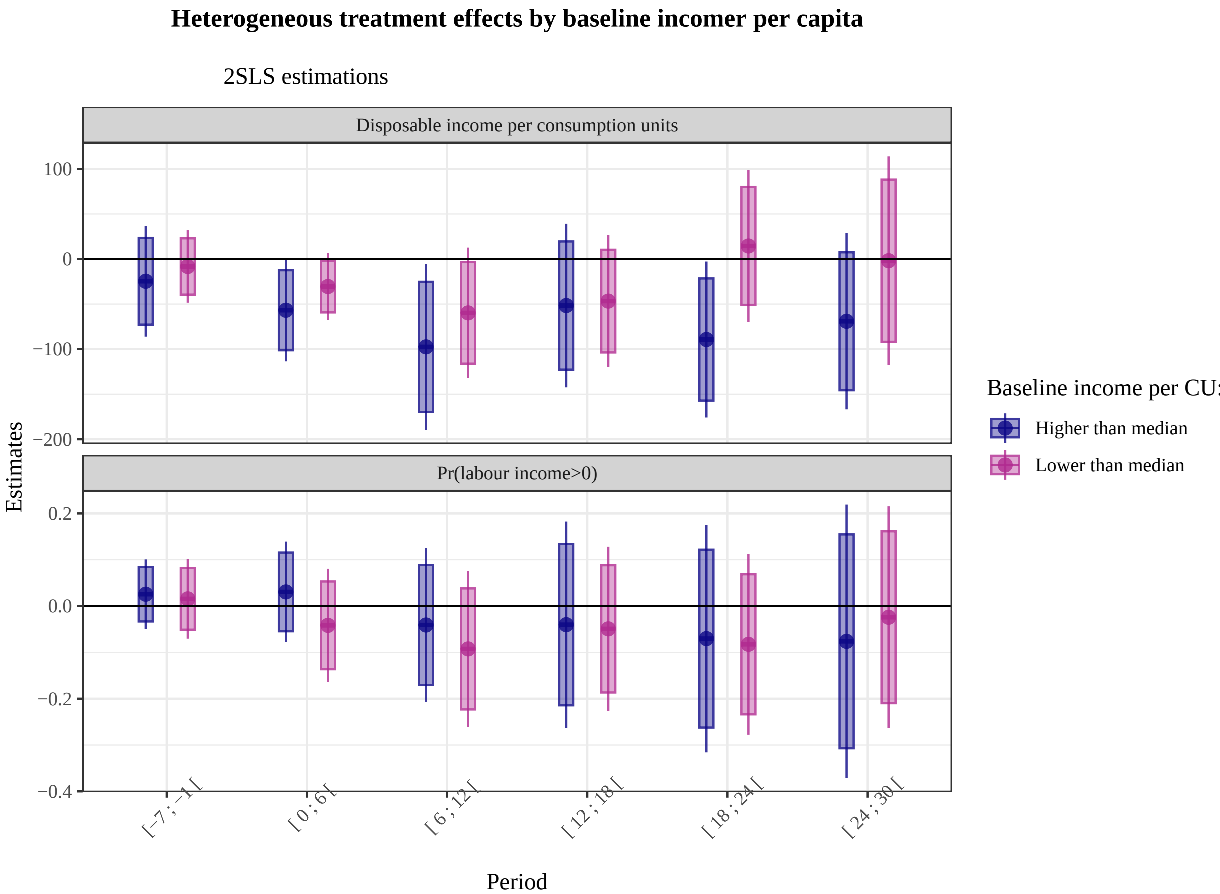
2SLS estimations



Sources: ALLSTAT 2017−01−01 to 2023−06−01 cohorts 2018 to 2021.
Notes: Estimations using TSLS with seniority x participation x period dummies
instrumented by the demeaned instrument x seniority x period dummies.
All models include blocks x cohort x relative months fixed effects. Covariates are measured at the month
before random assignment and include baseline level, initial assessment, education, French citizenship,
quartiles of age, receiving child support or family support allowance, housing benefits, children under 3,
children 3 to 5, children older than 16, taxable income higher than median and quartile of baseline income per capita.
All covariates are interacted with relative time dummies to have specific effects for each period.
Standard errors are cluster−heteroskedasticity robusts adjusted at the block x cohort level.
− Left panel: The dependent variable is the monthly disposable income per capita.
− right panel: The dependent variable is a dummy for positive labour incomes.
− Error bars indicate 95 % pointwise confidence intervals and extended lines account for FWER by subgroup.

D.VI Treatment effect heterogeneity by baseline income per capita

Figure D.27: Heterogenous treatment effects by social workers' initial assessment

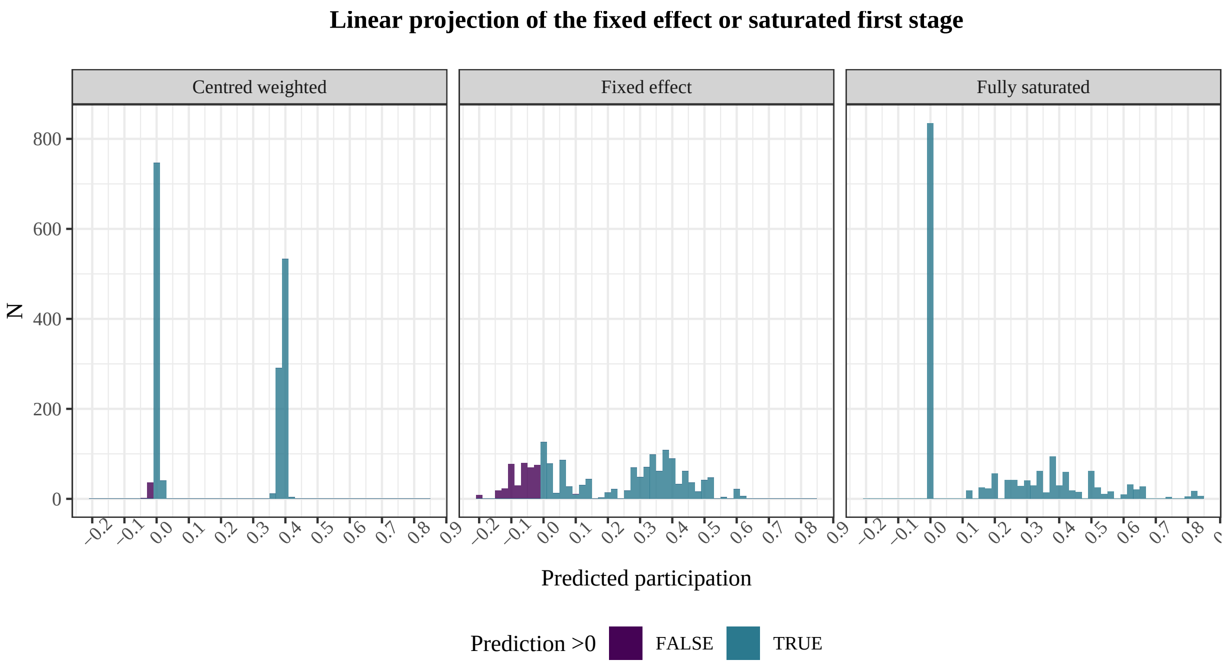


Sources: ALLSTAT 2017−01−01 to 2023−06−01 cohorts 2018 to 2021.
Notes: Estimations using TSLS with median income x participation x period dummies
instrumented by the demeaned instrument x median income x period dummies.
All models include blocks x cohort x relative months fixed effects. Covariates are measured at the month
before random assignment and include baseline level, initial assessment, education, French citizenship,
quartiles of age, receiving child support or family support allowance, housing benefits, children under 3,
children 3 to 5, children older than 16, taxable income higher than median and quartile of baseline income per capita.
All covariates are interacted with relative time dummies to have specific effects for each period.
Standard errors are cluster−heteroskedasticity robusts adjusted at the block x cohort level.
− Left panel: The dependent variable is the monthly disposable income per capita.
− right panel: The dependent variable is a dummy for positive labour incomes.
− Error bars indicate 95 % pointwise confidence intervals and extended lines account for FWER by subgroup.

E Robustness checks

E.I Predicted participation from the average first stage

Figure E.28: First stage predictions of the Fixed effect and saturated regressions

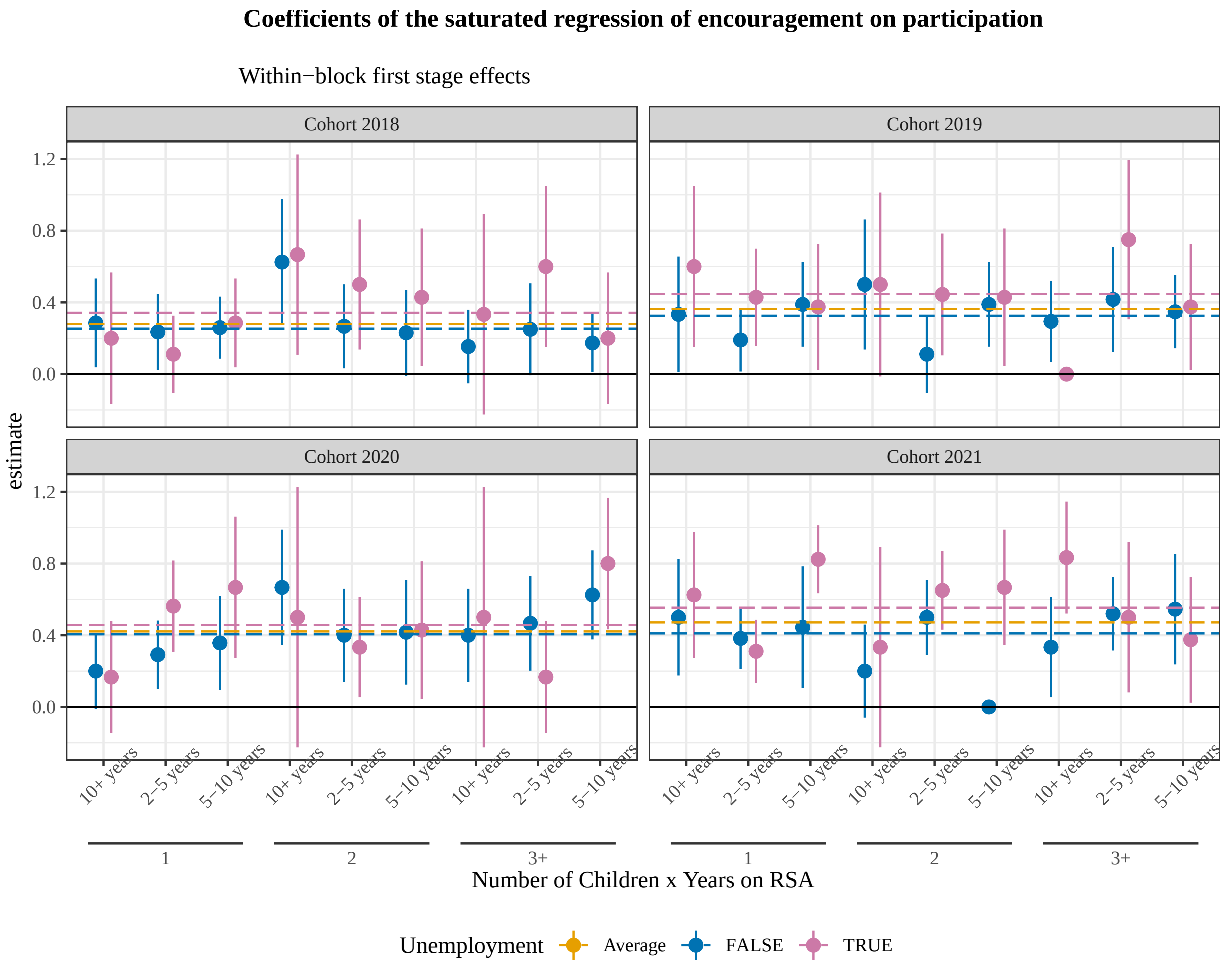


Sources: ALLSTAT, cohorts 2018 to 2021 at baseline.
Notes: Hisograms of the predicted values of the first stage. Binwidth of .02.
All models use inverse instrument propensity score weighting.
The fixed effect model regress participation on encouragement and blocks x cohorts FE.
The saturated model regress participation on the interactions between encouragement
and blocks x cohorts fixed effects.
The Centred weighted model remove block−specific propensity score from the treatment variable.

E.II Coefficients of the fully saturated first stage

In Figure E.29 I plot the coefficients and confidence intervals of the fully saturated first-stage regression; that is, the results of the regression of participation on encouragement × block × cohort fixed effects and block × cohort fixed effects. As this regression is fully saturated with only dummy variables, it estimates the conditional expectation function perfectly, and the coefficient of the interactions estimates the average effect of encouragement on participation in each block.

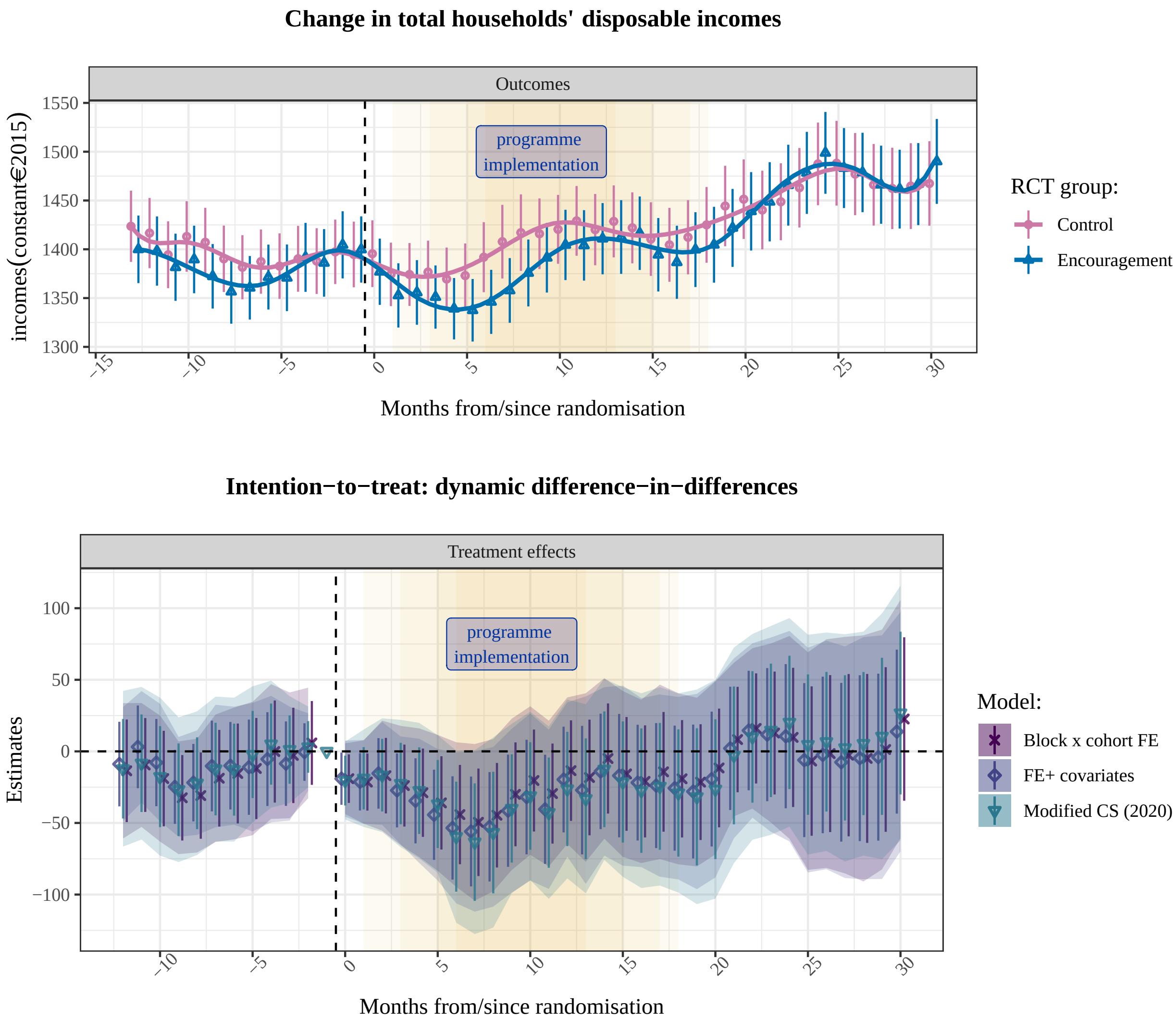
Figure E.29: First stage effect of encouragement on participation in each block



Sources: ALLSTAT cohorts 2018 to 2021 at the month of randomisation.
Notes : Coefficients and 95% confidence intervals of the regression of participation on fixed effects by strata x cohorts and the interaction of encouragement with these fixed effects.
Dashed lines indicate the average effect for each cohort and separate estimates based on unemployment registration.
Error bars are calculated using HC2 heteroskedasticity−robust standard errors.

A) Effects on total household incomes

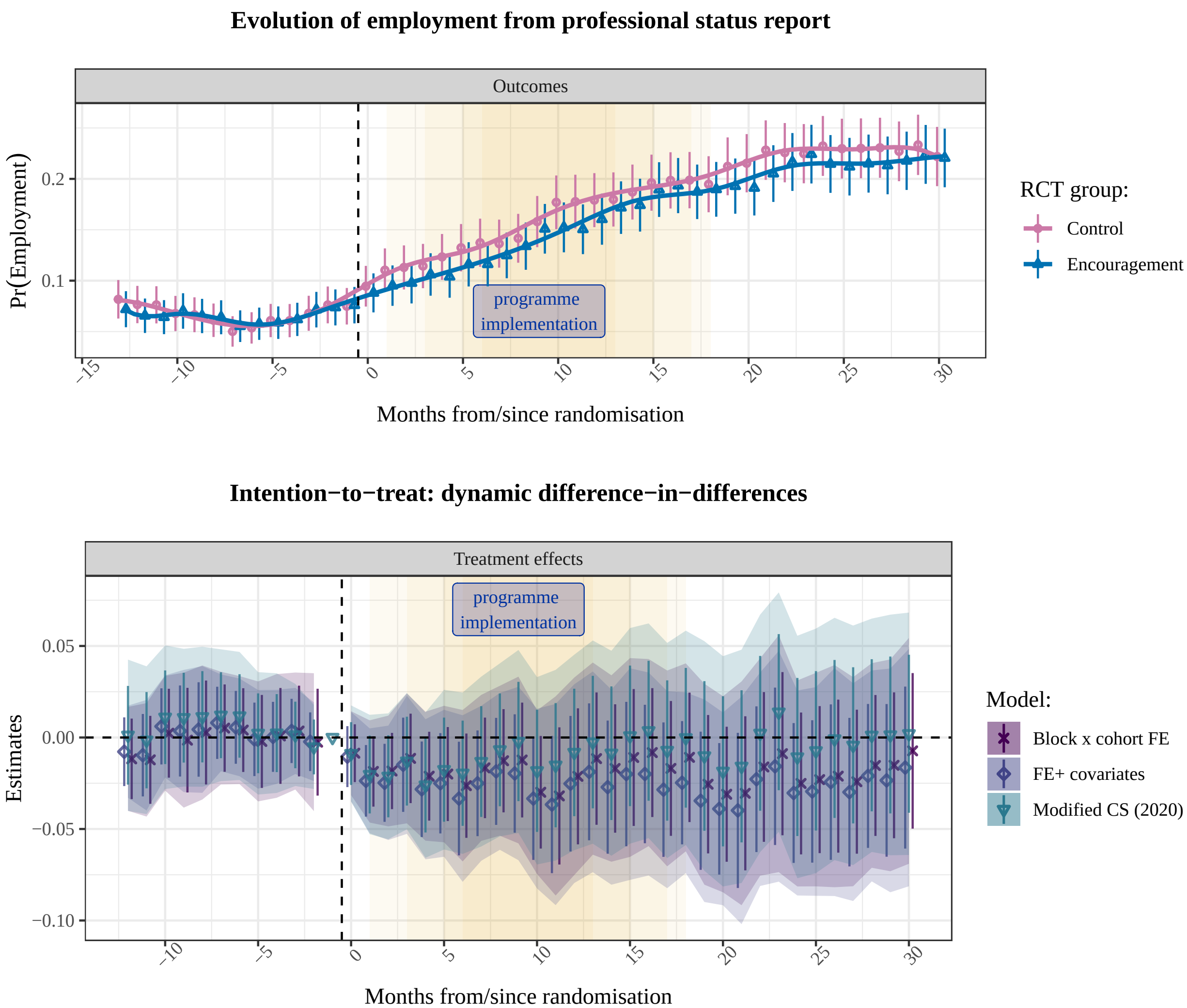
Figure E.30: Intention-to-treat effects of the programme on household disposable incomes estimated by difference-in-differences



Sources: ALLSTAT 2017−01−01 to 2023−06−01
Notes: The dependent variable is a dummy for positive individual labour income.
Top panel:
− Points indicate simple means by encouragement status over cohorts 2018 to 2021 in relative time since randomisation.
− Error bars indicate pointwise 95% CI using simple standard errors.
− Lines are conditional means estimated with spline regressions.
Bottom panel:
− Event study with block x cohort fixed effects with/without covariates using OLS, demeaned encouragement and inverse propensity score weighting or modified Callaway Sant' Anna (2020) estimator (See section IV).
− Error bars, indicates 95 % CI using cluster−robust standard errors at the block x cohort level.
− Shades indicates 95% CI adjusting for the FWER using the Holm−Bonferroni correction for OLS models and wild cluster bootstrap for modified Callaway and Sant' Anna (2020).
Covariates are measured at the month before random assignment and include baseline level, initial assessment, education, French citizenship, quartiles of age, receiving child support or family support allowance, housing benefits, children under 3, children 3 to 5, children older than 16, taxable income higher than median and quartile of baseline income per capita.
All covariates are interacted with relative time dummies to have specific effects for each period.
For OLS models, covariates also include dummies for being resampled in the 2022 cohort and being encouraged, interacted with relative time dummies. For the modified CS model, observations of the late encouraged group are dropped.

E.III Alternate measure of employment

Figure E.31: Intention-to-treat effects of the programme on household disposable incomes estimated by difference-in-differences

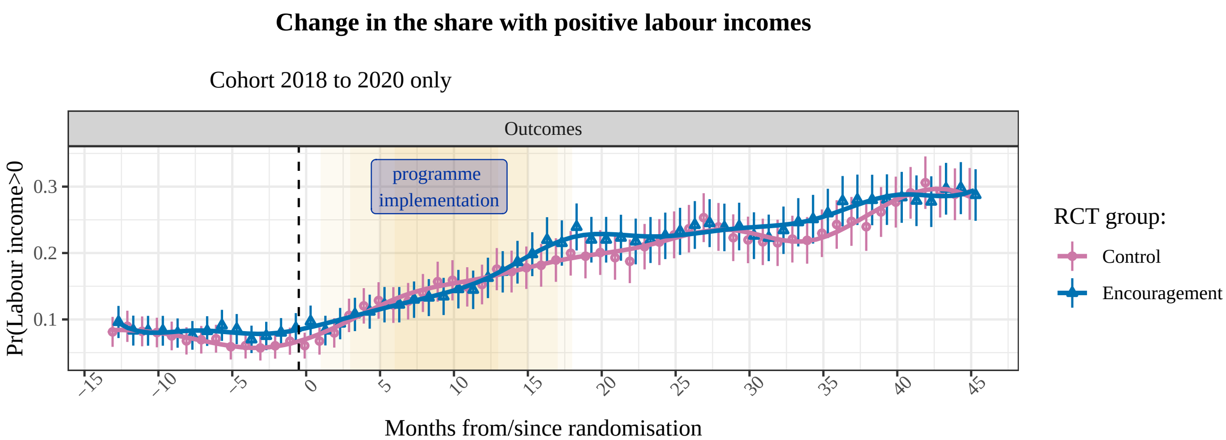


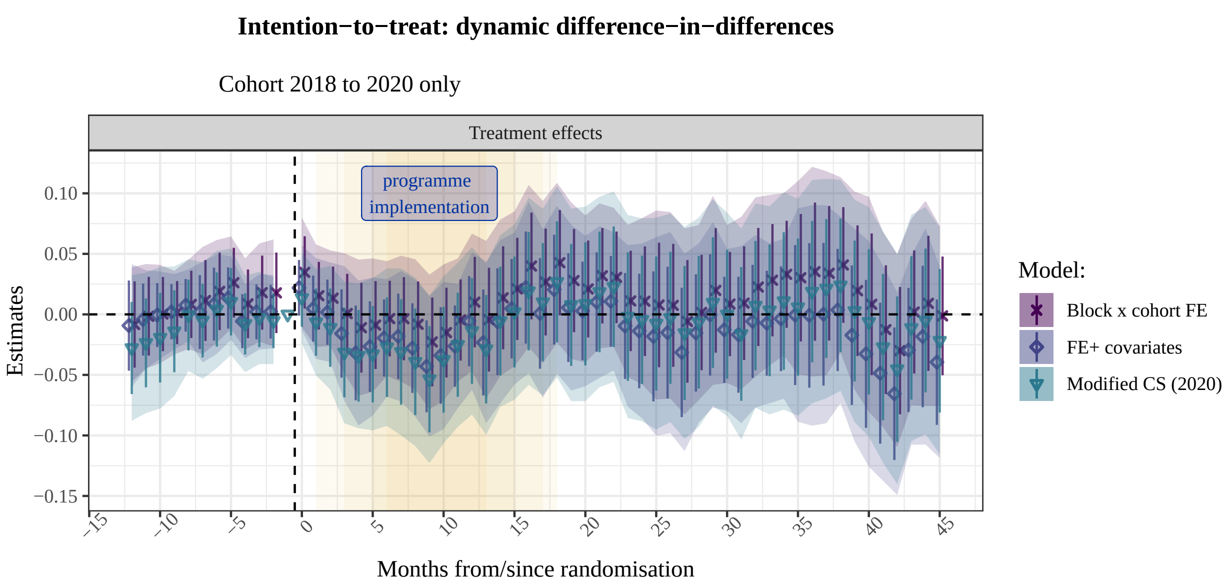
Sources: ALLSTAT 2017−01−01 to 2023−06−01
Notes: The dependent variable is a dummy for being classified as 'Active' in the professional status measure.
This variable is of poor quality as it filled by CAF Agents and ther may be delays.
Top panel:
− Points indicate simple means by encouragement status over cohorts 2018 to 2021 in relative time since randomisation.
− Error bars indicate pointwise 95% CI using simple standard errors.
− Lines are conditional means estimated with spline regressions.
Bottom panel:
− Event study with block x cohort fixed effects with/without covariates using OLS, demeaned encouragement and inverse propensity score weighting or modified Callaway Sant' Anna (2020) estimator (See section IV).
− Error bars, indicates 95 % CI using cluster−robust standard errors at the block x cohort level.
− Shades indicates 95%CI adjusting for the FWER using Holm−Bonferroni correction for OLS and wild cluster bootstrap for modified Callaway and Sant' Anna (2020).
Covariates are measured at the month before random assignment and include baseline level, initial assessment, education, French citizenship, quartiles of age, receiving child support or family support allowance, housing benefits, children under 3, children 3 to 5, children older than 16, taxable income higher than median and quartile of baseline income per capita.
All covariates are interacted with relative time dummies to have specific effects for each period.
For OLS models, covariates also include dummies for being resampled in the 2022 cohort and being encouraged, interacted with relative time dummies. For the modified CS model, observations of the late encouraged group are dropped.

E.IV Estimations for the three first cohorts up to 45 months

A) Intention-to-treat for cohort 2018 to 2020 on labour market participation

Figure E.32: Intention-to-treat effects of the programme on the labour market participation for the three first cohorts up to 45 months.

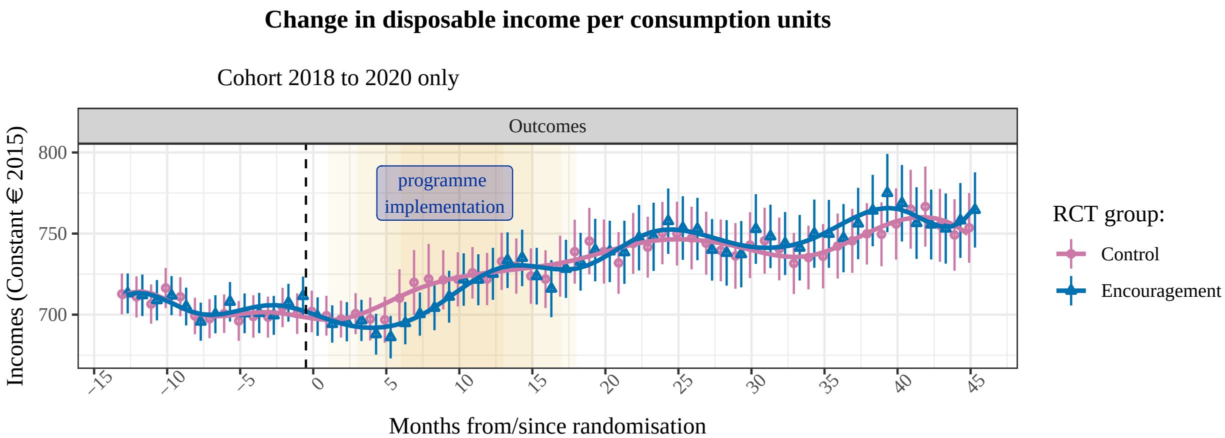


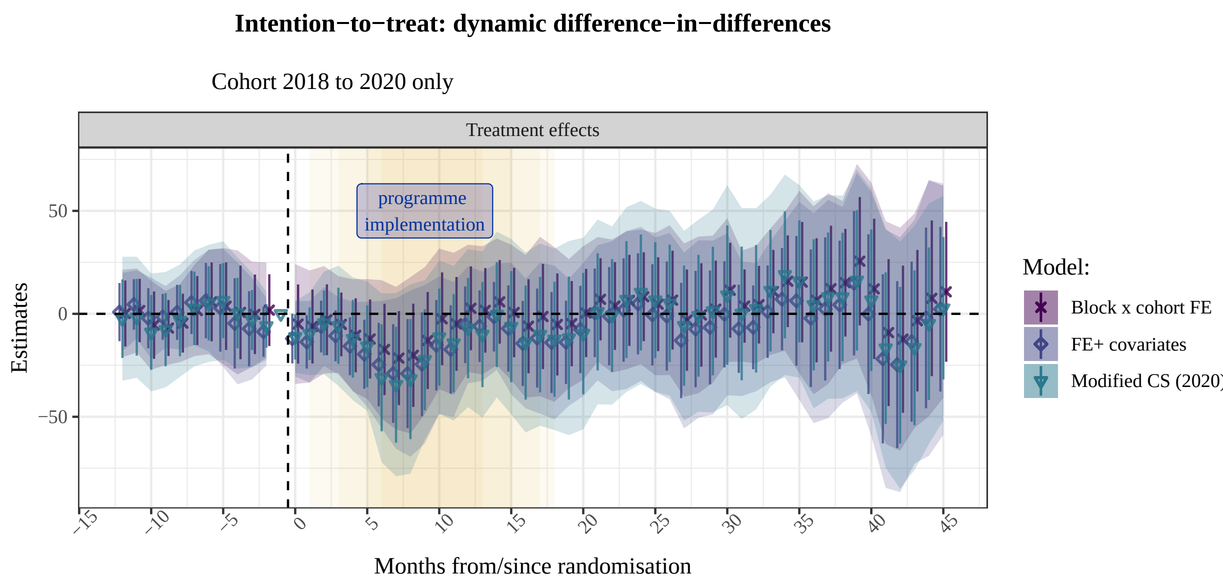


Sources: ALLSTAT 2017−01−01 to 2023−06−01
Notes: The dependent variable is a dummy for reporting positive labour income.
Top panel:
− Points indicate simple means by encouragement status over cohorts 2018 to 2020 in relative time since randomisation.
− Error bars indicate pointwise 95% CI using simple standard errors.
− Lines are conditional means estimated with spline regressions.
Bottom panel:
− Event study with block x cohort fixed effects with/without covariates using OLS, demeaned encouragement and inverse propensity score weighting or modified Callaway Sant' Anna (2020) estimator (See section IV).
− Error bars, indicates 95 % CI using cluster−robust standard errors at the block x cohort level.
− Shades indicates 95% CI adjusting for the FWER using Holm−Bonferroni correction for OLS and wild cluster bootstrap for modified Callaway and Sant' Anna (2020).
Covariates are measured at the month before random assignment and include baseline level, initial assessment, education, French citizenship, quartiles of age, receiving child support or family support allowance, housing benefits, children under 3, children 3 to 5, children older than 16, taxable income higher than median and quartile of baseline income per capita.
All covariates are interacted with relative time dummies to have specific effects for each period.
For OLS models, covariates also include dummies for being resampled in the 2022 cohort and being encouraged, interacted with relative time dummies. For the modified CS model, observations of the late encouraged group are dropped.

B) Intention-to-treat for cohort 2018 to 2020 on disposable income per capita

Figure E.33: Intention-to-treat effects of the programme on disposable income per capita for the three first cohorts up to 45 months.

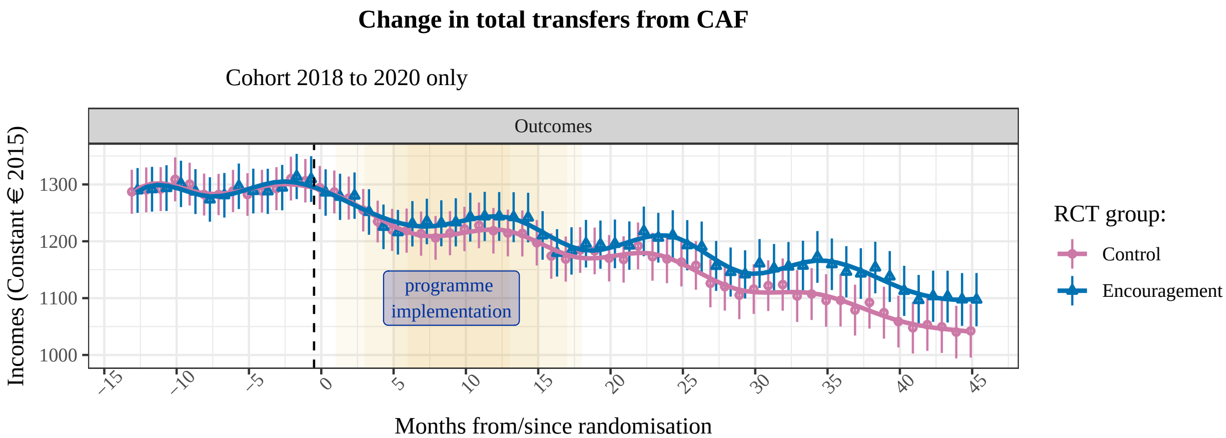


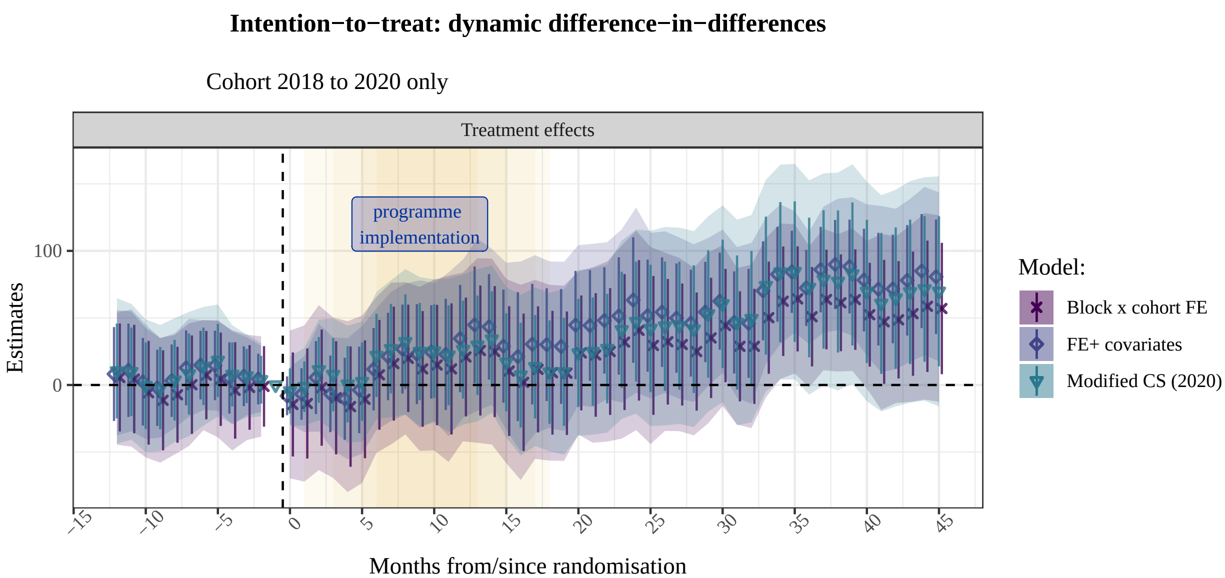


Sources: ALLSTAT 2017−01−01 to 2023−06−01
Notes: The dependent variable is the monthly disposable income per consumption units, in 2015 euros.
Top panel:
− Points indicate simple means by encouragement status over cohorts 2018 to 2020 in relative time since randomisation.
− Error bars indicate pointwise 95% CI using simple standard errors.
− Lines are conditional means estimated with spline regressions.
Bottom panel:
− Event study with block x cohort fixed effects with/without covariates using OLS, demeaned encouragement and inverse propensity score weighting or modified Callaway Sant' Anna (2020) estimator (See section IV).
− Error bars, indicates 95 % CI using cluster−robust standard errors at the block x cohort level.
− Shades indicates 95%CI adjusting for the FWER using Holm−Bonferroni correction for OLS and wild cluster bootstrap for modified Callaway and Sant' Anna (2020).
Covariates are measured at the month before random assignment and include baseline level, initial assessment, education, French citizenship, quartiles of age, receiving child support or family support allowance, housing benefits, children under 3, children 3 to 5, children older than 16, taxable income higher than median and quartile of baseline income per capita.
All covariates are interacted with relative time dummies to have specific effects for each period.
For OLS models, covariates also include dummies for being resampled in the 2022 cohort and being encouraged, interacted with relative time dummies. For the modified CS model, observations of the late encouraged group are dropped.

C) Intention-to-treat for cohort 2018 to 2020 on total transfers

Figure E.34: Intention-to-treat effects of the programme on total cash transfers for the three first cohorts up to 45 months.

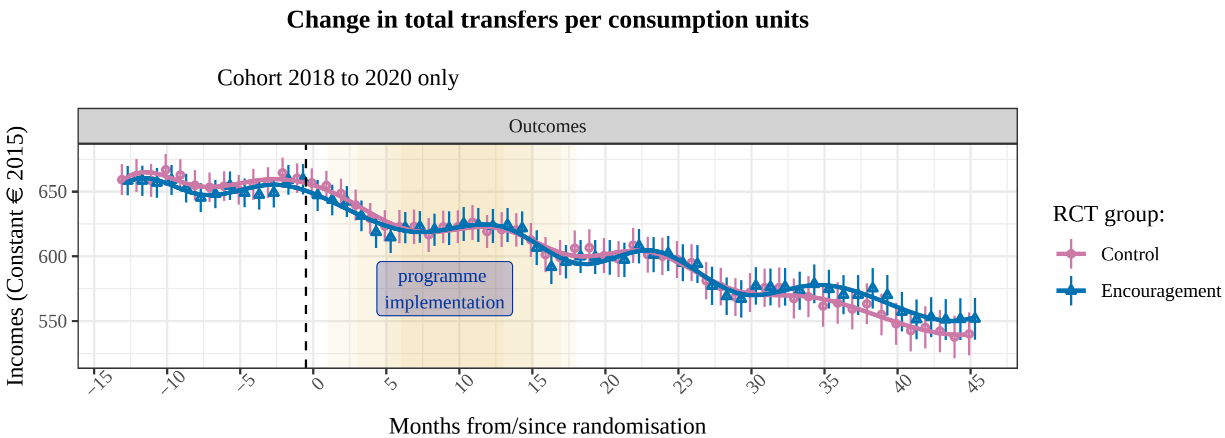


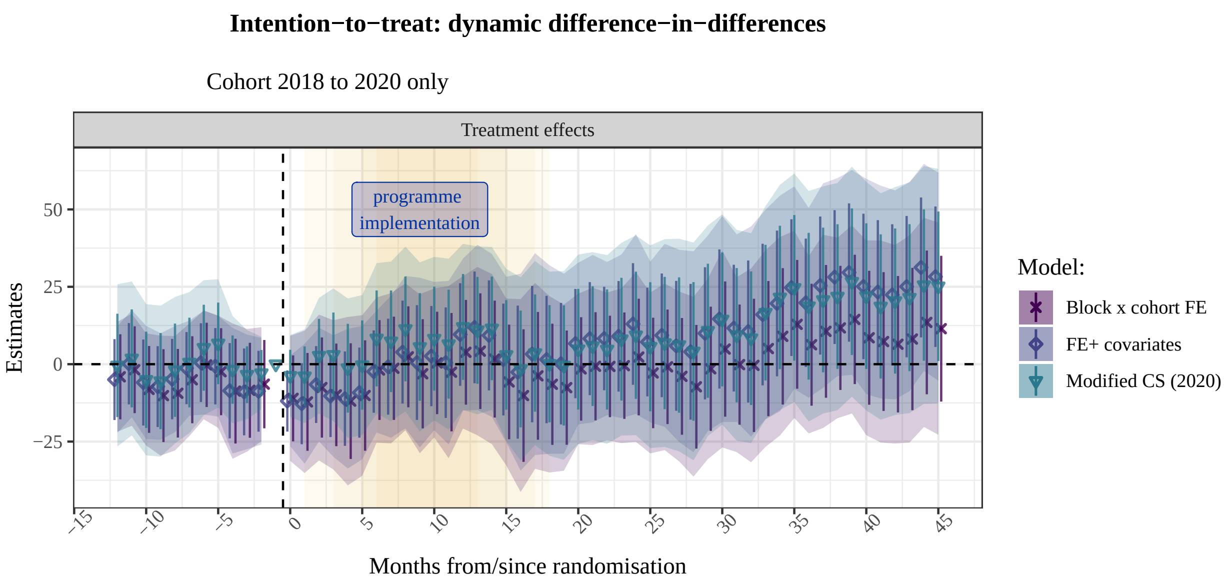


Sources: ALLSTAT 2017−01−01 to 2023−06−01
Notes: The dependent variable is the monthly total allowance paid by CAF, in 2015 euros.
Top panel:
− Points indicate simple means by encouragement status over cohorts 2018 to 2020 in relative time since randomisation.
− Error bars indicate pointwise 95% CI using simple standard errors.
− Lines are conditional means estimated with spline regressions.
Bottom panel:
− Event study with block x cohort fixed effects with/without covariates using OLS, demeaned encouragement and inverse propensity score weighting or modified Callaway Sant' Anna (2020) estimator (See section IV).
− Error bars, indicates 95 % CI using cluster−robust standard errors at the block x cohort level.
− Shades indicates 95%CI adjusting for the FWER using Holm−Bonferroni correction for OLS and wild cluster bootstrap for modified Callaway and Sant' Anna (2020).
Covariates are measured at the month before random assignment and include baseline level, initial assessment, education, French citizenship, quartiles of age, receiving child support or family support allowance, housing benefits, children under 3, children 3 to 5, children older than 16, taxable income higher than median and quartile of baseline income per capita.
All covariates are interacted with relative time dummies to have specific effects for each period.
For OLS models, covariates also include dummies for being resampled in the 2022 cohort and being encouraged, interacted with relative time dummies. For the modified CS model, observations of the late encouraged group are dropped.

D) Intention-to-treat for cohort 2018 to 2020 on total transfers per consumption units

Figure E.35: Intention-to-treat effects of the programme on total cash transfers per consumption units for the three first cohorts up to 45 months.





Sources: ALLSTAT 2017−01−01 to 2023−06−01
Notes: The dependent variable is the monthly total allowance paid by CAF in 2015 euros divided by the number of consumption units.
Top panel:
− Points indicate simple means by encouragement status over cohorts 2018 to 2020 in relative time since randomisation.
− Error bars indicate pointwise 95% CI using simple standard errors.
− Lines are conditional means estimated with spline regressions.
Bottom panel:
− Event study with block x cohort fixed effects with/without covariates using OLS, demeaned encouragement and inverse propensity score weighting or modified Callaway Sant' Anna (2020) estimator (See section IV).
− Error bars, indicates 95 % CI using cluster−robust standard errors at the block x cohort level.
− Shades indicates 95%CI adjusting for the FWER using Holm−Bonferroni correction for OLS and wild cluster bootstrap for modified Callaway and Sant' Anna (2020).
Covariates are measured at the month before random assignment and include baseline level, initial assessment, education, French citizenship, quartiles of age, receiving child support or family support allowance, housing benefits, children under 3, children 3 to 5, children older than 16, taxable income higher than median and quartile of baseline income per capita.
All covariates are interacted with relative time dummies to have specific effects for each period.
For OLS models, covariates also include dummies for being resampled in the 2022 cohort and being encouraged, interacted with relative time dummies. For the modified CS model, observations of the late encouraged group are dropped.

Chapter 3.

Tax burden on the poor: Single mothers' optimisation behaviours following an experimental activation programme in France

« *'Faites des enfants c'est fantastique vous vous sentirez plus femmes et accomplies que jamais', mais faites-les dans une société en dégringolade, où le travail salarié est une condition de survie sociale, mais n'est garanti pour personne, et surtout pas pour les femmes. Enfantez dans des villes où le logement est précaire, où l'école démissionne, où les enfants sont soumis aux agressions mentales, les plus vicieuses, via la pub, la télé, internet, les marchands de sodas et confrères. Sans enfant, pas de bonheur féminin, mais élever des gamins dans des conditions décentes sera quasi impossible. Il faut, de toutes façons, que les femmes se sentent en échec* ».

Virginie Despentes (2006), King Kong Theorie, Grasset. Paris.

This article is part of Arthur Heim's PhD dissertation. He wants to thank Marc Gurgand for his supervision and guidance over the years and Karen Macours for her advice and support. Special thanks to Jules Cornetet and Quynh-Chi Doan for their insights on the results of microsimulation part of the research, Clemence Helfter and Saad Loutfi for numerous comments and stimulating discussions. We are grateful to Bruno Palier, Antoine Bozio, Robin Huguenot-Noël, Michael Zemmour, Clement Carbonnier and the participants of the Labour and public policies seminar of Paris School of Economics, the Séminaire Travail en Économie Politique (STEP) of Paris I Panthéon Sorbonne and the informal research group on poverty at LIEPP (Sciences Po) for their stimulating comments, questions and suggestions at various steps of this project.

Abstract

In this paper, we challenge the idea that the French tax-benefit system ”makes work pay” for single parents and analyse reactions following a randomised intensive welfare-to-work programme rolled-out from 2018 to 2022. The 2019 reform of in-work benefits was adopted in the timeline of this experiment. The intervention targetted single parents on long-term welfare and directly provided individualised and detailed information on rights and benefits, in a year-long activation programme likely to have further reduced psychological barriers and search costs. We use this experiment to measure the reactions of financially disadvantaged single parents to the reformed tax-benefit system. By employing open-source models, we demonstrate that the composition of total social transfers varies across household compositions, altering incentives. Additionally, our simulations reveal that the implicit marginal tax rate for single parents exceeds that of couples across all scenarios we explored, with a range between 50% and 60% of the full-time minimum wage representing the lowest implicit tax rate, irrespective of household structure. It is bounded above by an implicit marginal tax rate roughly twice as high. Beyond full-time minimum wages, single parents with one or two children face an implicit marginal tax rate exceeding 70%. Once single parents accept welfare provision, they unknowingly sign-in for the highest tax burden of the income distribution. We coined the term ”*Assistaxation*” to convey the idea of providing assistance in a way that becomes burdensome, overly taxing – either mentally, physically, emotionally, or financially – and very hard to escape. . Our primary contribution lies in leveraging experimental variations in assignment probabilities to infer the counterfactual distribution of untreated compliers within an instrumental variable framework, using semi-parametric weighted distribution regressions. These estimates of counterfactual densities enable us to measure the bunching mass at kink points and estimate the elasticity of labour income for treated compliers. We find substantial elasticities of labour income, approximately two, which are ten times higher than those of the comparison group. Furthermore, our analysis indicates that job re-entry leads to lower growth of disposable income for treated compliers compared to untreated compliers, resulting in increased in-work poverty. Lastly, our examination of the programme’s effects on family structure reveals significant heterogeneous effects based on the number of children at baseline. Our findings highlight the considerable tax burden faced by impoverished single-parent households, acting as a significant disincentive to employment for some, while encouraging part-time employment more generally. In either scenario, these incentives perpetuate situations in which households lack sufficient income to escape poverty, thereby leading to reliance on high levels of social transfers.

- **JEL Classification Numbers** : I38, J16, J18
- **Keywords**: Welfare-to-Work, single parents, active labour market policy, non-parametric models, distribution regressions
- **Authors’ contribution**: Arthur Heim was responsible for the experimental design, building datasets and econometic analysis. Alexandra Galitzine provided critical feedback and helped shape the research, analysis and manuscript. Both Alexandra Galitzine and Arthur Heim contributed to the final version of the manuscript.

Résumé

Dans cet article, nous remettons en question le narratif selon lequel le système socio-fiscal français "rend le travail payant" pour les familles monoparentales et analysons les réactions suite à un programme randomisé intensif d'accompagnement social et professionel déployé de 2018 à 2022. Ainsi, la réforme de 2019 de la prime d'activité a été adoptée dans la période d'évaluation. L'intervention ciblait des familles monoparentales au RSA depuis plusieurs années et leur a directement fourni des informations individualisées et détaillées dans le cadre d'un accompagnement global d'une durée d'un an, susceptible d'avoir réduit, davantage encore, les divers freins à l'emploi. Nous utilisons cette expérience pour mesurer les réactions des familles monoparentales financièrement défavorisées au système socio-fiscal nouvellement réformé. En utilisant des modèles *open source* de ce dernier, nous démontrons que la composition des transferts sociaux totaux varie selon les compositions des ménages, modifiant fortement les incitations. De plus, nos simulations révèlent que le taux marginal implicite d'imposition pour les familles monoparentales dépasse celui des couples dans tous les scénarios que nous avons explorés, avec une fourchette entre 50% et 60% du SMIC à temps plein représentant le taux d'imposition implicite le plus bas, quel que soit la structure du ménage. Cependant, celui-ci est borné par un taux marginal d'imposition implicite environ deux fois plus élevé. Au-delà du salaire minimum à temps plein, les familles monoparentales avec un ou deux enfants font face à un taux marginal d'imposition implicite dépassant 70%. Sans le savoir, celles qui demandent le RSA signent en même temps pour le plus haut taux de taxation de toute la distribution de revenu. Nous avons proposé le terme "*Assistaxation*" pour désigner ce phénomène consistant à taxer massivement les ressources économiques, physiques et mentales des personnes ayant recours à l'aide publique, leur laissant au passage peu de moyen de s'en extraire. Notre principale contribution réside dans l'exploitation des variations expérimentales des probabilités d'affectation pour déduire la distribution contrefactuelle par variable instrumentale, en utilisant des régressions de distribution semi-paramétriques pondérées. Nous utilisons ces estimations des densités contrefactuelles pour estimer la masse de revenus aux coudes des contraintes budgétaires et mesurer l'élasticité du revenu du travail des participantes. Nous obtenons des élasticités élevées d'environ 2, soit 10 fois que celle du groupe de comparaison. De plus, notre analyse indique que la reprise d'emploi augmente moins le revenu des participantes que dans le contrefactuel, entraînant une augmentation de la pauvreté laborieuse. Enfin, notre analyse des effets du programme sur la structure familiale révèle des effets hétérogènes significatifs en fonction du nombre d'enfants au début de l'étude. Nos résultats soulignent le fardeau fiscal considérable auquel sont confrontées les familles monoparentales défavorisées, qui constitue un puissant frein à une reprise d'activité pour certaines, tout en incitant fortement à l'emploi à temps partiel pour celles qui auraient travaillé davantage. Dans tous les cas, ces incitations perpétuent des situations où les ménages ne disposent pas de revenus suffisants pour sortir de la pauvreté, tout en maitenant une forte dépendance aux aides sociales.

- **Codes Journal of economic literature** : I38, J16, J18
- **Mots clés**: Welfare-to-Work, single parents, active labour market policy, non-parametric models, distribution regressions
- **Contribution des auteurs et autrices** : Arthur Heim a été responsable du design expérimental, de la préparation et de l'analyse des données. Alexandra Galitzine a fourni des commentaires critiques et a contribué à façonner la recherche, l'analyse et le manuscrit. Alexandra Galitzine et Arthur Heim ont tous deux contribué à la version finale du manuscrit.

I Introduction

The debate surrounding redistribution policies through progressive taxation and generous transfers revolves around the tension between social justice and economic efficiency. On the one hand, redistribution is essential for ensuring the economic well-being of the less fortunate, acknowledging that differences in earnings often stem from factors beyond individuals' control, such as innate ability, social background or sheer (bad) luck (Fleurbaey and Maniquet 2018). On the other hand redistributive measures can hamper incentives to work among both the wealthy and recipients of transfers, potentially undermining overall economic productivity (Ebert 2005; Piketty and Saez 2013). Over the past three decades, traditional Welfare States have been put under the pressure of fiscal consolidation and shifted towards active labour market policies (ALMP) (J. P. Martin 2015; Crépon and van den Berg 2016a). These welfare reforms took various forms around the world but dramatically reshaped labour market institutions and social insurance mechanisms with welfare-to-work programmes, in-work benefits, payroll tax exemptions, training initiatives, job search assistance and monitoring and so on. The US Earned Income Tax Credit (EITC) plays a special role in the economics of welfare reforms. It has been extended and inspired many similar policies around the world, although there is an ongoing debate among academics on its effects on employment (See Section II).

France's adoption of the *Revenu de solidarité active* (RSA) in 2008 is one of such examples and subsequent reforms in 2016 and 2019 – separating the in-work benefit into *Prime d'activité* (PA) and increasing the amounts higher up the wage distribution – underscore the ongoing efforts to incentivise work and increase disposable incomes of low wage workers (Gurgand and Margolis 2008; Bargain and Doorley 2011; Bargain and Vicard 2014; Sicsic 2019).

However, these policies rely on several pre-requisites. First, people need to understand how benefits change their incentives to work. When people have limited understanding of the actual tax-benefit schedules they face, they are likely to perceive them in a crude fashion and under-react to the incentives. Recent work shows that households are often unaware or misinformed about the incentives of tax-benefit systems. In particular, they have problems understanding non-linearities in the system and rely instead on some "*mentally linearised*" tax rate (Rees-Jones and Taubinsky 2020; Caldwell, Nelson, and Waldinger 2023). Uncertainty may be the result of more complex features of the tax code, such as the phase-in and phase-out regions for tax-based transfer programmes or rules for married tax filers. Conversely, providing additional information can trigger large reactions at the intensive margins, inducing large distributional effects (Raj Chetty, Friedman, and Saez 2013; Raj Chetty and Saez 2013; Kostøl and Myhre 2021). However, the means by which information are provided seem to be equally important; many information-only and nudge experiments have proven ineffective (Linos et al. 2020; Nyman, Aggeborn, and Ahlskog 2023) or delivered much lower effects than human assistance (Castell et al. 2022; Finkelstein and Notowidigdo 2019; Bergman et al. 2019).

Second, another prerequisite is that there should actually be incentives to work. The latter may seem obvious considering the amount of public money spent in active labour market policies. However, it may not be the case for everyone, and there may be large frictions impeding adjustments. At the macro-economic level, many scholars underlined the paradox that despite massive increases in public spending in active labour market policies, long-term unemployment and poverty have not been reduced (Vandenbroucke and Vleminckx 2011a; Jaehrling, Kalina, and Mesaros 2015; Van Winkle and Struffolino 2018). At the micro-level, a large body of evidence shows that active labour market policies have little effects on labour market outcomes of single parents and often detrimental effects on their health and well-being (Cook 2012; Pega et al. 2013; Gibson et al. 2018), sometime extending to their children (Løken, Lommerud, and Holm Reiso 2018).

In this paper, we challenge the idea that the French tax-benefit system "makes work pay" for single parents and analyse reactions following an intensive welfare-to-work programme of single parents on long-term welfare. The intervention called **Reliance** was rolled-out each year from 2018 to 2022 using a randomised encouragement design to recruit each cohort. The sample includes 1666 households from the 4 first cohorts; 826 controls and 840 single parents in the encouragement groups. 95% of the sample are single mothers followed through for at least 30 months after random assignment. Our data come from the National family allowance fund (Cnaf), the administration in charge of welfare payment, and include all social transfers and reported incomes, allowing to measure labour and non-labour incomes by spouse. This high stake programme has drawn 4 times more resources than the usual budget for social support of welfare recipients to offer a year-long programme to 82 participants each year on average (mean take-up $\approx$ 39%). The intervention was designed to provide both individual and group support through thematic workshops, sometimes involving other institutions. The primary focus of the intervention was to create and validate realistic professional projects, aligning expectations with job opportunities and enhancing job search efficiency. Additionally, the intervention provided extensive social support related to parenthood, self-esteem, gender norms, and other relevant topics. Furthermore, the intervention included several workshops on accessing rights and direct support to alleviate the emotional burden associated with administrative procedures. Notably, regular sessions with social workers from the Family allowance fund (Caf) were organized to help participants understand and access their rights, and we also provided simplified plots of the amounts of social benefits over incomes from simulation models of the Family allowance fund.

Our previous evaluation showed no average effect of the programme on employment and disposable income beyond an expected lock-in effect but revealed heterogeneous treatment effects, particularly concerning employment and disposable income by number of children at baseline (Heim 2024). The key message of this paper is that selection effects are important and hard to neutralise without random assignment. Considering the lack of evidence on similar programmes in Europe, this paper focused on clean identification on outcomes typically investigated in labour economics. Drawing on additional insights from the qualitative evaluation, our primary hypothesis is that while preparing participants for future job re-entry, the programme enhanced their knowledge and understanding of the tax-benefit system. With intensive social support and tailored interventions, it could provide them with the means to define and pursue achievable goals with a clearer understanding of the consequences of their choices.

In general, observing the effects of information requires participants to actually learn things, so incentives need to be largely unknown prior to the intervention. They also need to be sufficiently large to induce changes, and adjustments must be fairly quick. Moreover, the causal mechanism must concern micro-level behaviour, as opposed to things like collective wage bargaining, for the identifying assumptions to be valid (Nyman, Aggeborn, and Ahlskog 2023).

In the timeline of this experiment, the 2019 reform of in-work benefits was adopted unexpectedly in the last quarter of the first cohort's training when the second was being recruited. It was followed by a large increase in the number of recipients, largely coming from previously ineligible households higher on the wage distribution, but also from newly registered (Dardier, Doan, and Lhermet 2022). However, current evaluations are unable to disentangle the extensive margin response from the reduction in non take-up[1].

Our setting ends-up being an ideal framework to measure the reactions of poor single parents to the reformed tax-benefit system. While we cannot measure the effect of the reform *per se*, as all cohorts have been exposed, we can measure if the programme had effects that are consistent with the incentives of the tax-benefit system. Our intervention directly provided individualised and detailed information in a year-long social support likely to have further reduced psychological barriers and search costs. If that is the case, our experiment reveals *quasi-frictionless* elasticities of labour for single parents on long-term welfare. In a way, we leverage this randomised experiment to uncover more structural reactions to the French tax-benefit system, which we therefore need to document. Our investigation is thus guided by two research questions:

1) *How does the tax-benefit system adjust to earnings and family composition ?*
2) *How do single parents react to the incentives of the tax-benefit system*?

We build our arguments in three steps. First, we review the literature on welfare reforms and single parents and build a simple theoretical framework grounded in the literature on bunching. The idea is to think of the way the programme may have affected participants introducing imperfect knowledge, psychological barriers, and adjustment costs with models commonly employed in recent similar research (R. Chetty et al. 2011; Henrik J. Kleven and Waseem 2013; Kostøl and Myhre 2021). We come back to the model with empirical estimates of bunching mass and use this simple framework to estimate and compare the observed elasticities across groups. However, our primary argument is a critique constructed from recent empirical findings, perspectives from feminist economics, and advancements in household economics. Second, we use an open-source model of the tax-benefit system to simulate social transfers by family structure and size, estimate implicit marginal tax rates and demonstrate the disparities and divergent incentives faced by single mothers with different numbers of children. Third, we analyse the causal effects of the programme on income distributions and family structures.

[1] Bozio et al. (2023) worked for France stratégie to evaluate the effect of the reform and we were part of the reviewing committee. Their identification strategy based on difference in difference by number of children was rejected after a placebo test on data from the year before. The report is inconclusive but did not remain in a file drawer, which we think is a good thing for science.

As an answer to our first research question, simulations allow us to illustrate three under-reported[2] stylised facts of the French benefit system: First, the composition of total social transfers is very different across household' structures. Specifically, the greater the number of children, the higher the family and housing benefits, while the lower the RSA and PA. Second, PA amounts are smaller and deplete at faster rates after the minimum wage for single mothers than for couples. Third, the implicit marginal tax rate of single parents is higher than couples in all the simulations we ran, but there is a range between 50% and 60% full-time minimum wage with minimal implicit tax rate and stable levels of all social transfers.

An important consequence of the first point is that different social transfers have different schedules and eligibility rules and are more or less sensitive to adjustments of family structure and incomes. For instance, housing benefits are the same amount with or without a partner, but use the same formula with one or two incomes. The complex interactions between various social transfers with differing schedules result in a much higher tax burden for single parents than for couples. That is mostly because RSA and PA are differential incomes and while their baseline amount increases with the number of children, they entirely deduce all family benefits, a lump-sum amount if they receive housing benefits, and all child support single parents may receive. That means that the higher their amounts of non-labour incomes - including payment from non-custodial parents - the lower welfare and in-work benefits they get. The family support allowance (ASF) – which substitutes child support under strict conditions – is also largely deduced from RSA and PA. As Pucci and Périvier (2022) noted before us, the French system saves welfare spending by taxing 100% of child support of single-parents with RSA or PA allowance. For single parents on welfare, receiving child support reduces their incomes if they received ASF, and changes nothing otherwise.

Ultimately, the range between 50% and 60% full-time minimum wage has the lowest implicit tax rate of the income distribution whatever the household structure, and is bounded by an implicit marginal tax rate roughly twice as high above. Beyond full-time minimum wage, single parents with one or two children face an implicit marginal tax rate higher than 70%. That means that the most typical labour contract for workers with little experience or education is the most heavily taxed of the distribution, making-it very hard for single parents to increase their income through work. In practice, a single mother with two children working full-time on minimum wage would only keep € 23 out of a € 100 net wage increase, with a reduction implemented from the following quarter. Couples very rarely have that level of tax rate anywhere in the income distribution. We discuss additional incentives in the paper, in particular regarding cohabitation and number of children.

Overall, the system is overwhelmingly complex. The precise incentives are not known, even to otherwise well-informed individuals[3]. This lack of salience hinders informed decision-making to the point where it wouldn't be far-fetched to talk about *obfuscation*, *i.e.* intentionally making something more complex or unclear to hide the truth, confuse others, or obscure information. While administrative burden already creates large costs for recipients, the complex interactions between different social transfers - with different schedules - inflates the cost of exit. Once single parents accept welfare provision, they unknowingly sign-in for the highest tax burden of the income distribution. We use the term *Assistaxation* to convey the idea of providing assistance in a way that becomes burdensome, overly taxing – either mentally, physically, emotionally, or financially – and very hard to escape.

Moving on to the data from this experiment, we present three main sets of estimations. First, we estimate typical bunching estimators comparing participants with non-participants using polynomial regressions on the density of observations around kink-points. Adjusting for diffuse bunching, round numbers and other kink points, we find that participants bunch at kink points, with few reporting incomes exceeding 75% of the full-time minimum wage. Facing the highest variations and levels of implicit tax rates, the bunching is particularly pronounced among single parents with two children at baseline. Conversely, the distributions of labour incomes among the control group and never-takers exhibit much lower and diffuse bunching mass at the 60% kink point and another mass at the full-time minimum wage. These first estimates show large differences between groups and also reveal that without the programme, single parents seems pretty un-reactive to the large variations in monetary incentives they face. Using our estimates of the implicit marginal tax rates around 60% minimum wage and bunching mass plugged-in an iso-elastic utility function, the observed elasticities are between .20 and .30 for non-participants. These range of estimates are close to those found in the literature. However, the same elasticity around kink-points among participants is closer to 1.

[2] The recent exceptions are the dedicated reports by Périvier (2022a) and Pucci and Périvier (2022), backing research papers (Allègre, Périvier, and Pucci 2021), and the book coordinated by Le Pape and Helfter (2023).

[3] Despite 6 years of PhD studying this system, we still uncover new rules on a regular basis.

Our main contribution leverages the experimental variations of assignment probabilities to infer the counterfactual distribution of untreated compliers in an instrumental variable framework using semi-parametric weighted distribution regressions. Our identification strategy is facilitated by one-sided non-compliance and a strong first stage allowing to identify the entire distribution of potential outcomes of treated and untreated compliers using causal weights (Frölich and Melly 2013). Unlike the κ-weighting representation of the LATE of A. Abadie (2003), for which some estimators can create negative weighting schemes (Słoczyński, Uysal, and Wooldridge 2022a), one-sided non-compliance ensures positive weights and an average-treatment-effect-on-the-treated (ATT) interpretation of instrumental variable estimates (Frölich and Melly 2013). The instrument propensity score is given by design with .5 in expectation but we use a Probit to predict the individual propensity score accounting for uneven numbers of individuals in blocks and imbalance in small blocks.

We then estimate the distribution of potential labour income of treated and untreated compliers using causal weights with the semi-parametric distribution regression estimator proposed by Matias D. Cattaneo, Jansson, and Ma (2021). The latter implements local polynomial regressions with data-driven optimal bandwidth and provide de-biased simultaneous confidence intervals around the estimated distributions. Since our estimates are based on strictly positive values of income, the densities of 0 income is missing and the estimates could neglect these extensive margin differences. However, this estimator also allows to rescale the distributions. We can therefore estimate the potential outcome masses at 0 using TSLS and use them to rescale the potential outcomes' densities. With this feature, we are able to visualise reactions at both the extensive and intensive margins simultaneously.

Our estimates confirm the sharp bunching and show that the distribution of incomes of untreated compliers would have been higher had they not participate. We use these estimates of the counterfactual densities to estimate the bunching mass at kink point and retrieve the elasticity of labour income of treated compliers. The latter is more than twice as high as the one estimated with typical bunching methods. We confirm these results with estimates of quantile intention-to-treat effects and an additional analysis on the differential effect of job re-entry for treated and untreated compliers using an instrumented triple difference. We leverage the variation in the timing of job re-entry and the share of treated compliers at each date to measure the change of disposable income for treated and untreated compliers in an event-study like estimation. These results confirms that job re-entry causes lower growth of disposable incomes for treated compliers relative to job re-entry for untreated compliers, and increases in-work poverty.

Finally, our analysis of the effects of the programme on family structure also reveals large heterogeneous effects by number of children of baseline. In brief, those with one child are more likely to re-partner, those with two children are less likely to get pregnant for the year following the programme while parents of three or more children are more likely to remain with their older children longer. These effects are large and lasting, showing that active labour market programmes affect many important decisions going far beyond labour market participation. Ultimately, our final result confirms the analysis of Heim (2024) showing a precise null effect on disposable income per consumption unit on the entire quantile distribution. While the programme affected many important decisions and induced large changes, these reactions end up leaving their disposable income at the same level it would have been without the programme, but with critical differences in their composition.

This second analysis allows to better understand the results at the extensive margin detailed in Heim (2024). The latter shows that the programme has no average effect but enrolls single mothers most likely to find a job on their own. These new results show that the programme increased their understanding of the tax-benefit system in a consequential way. As a result, compliers with two children work part-time instead of full-time; compliers with one child are less likely to work, more likely to report living with their partner, aligning the households' earned incomes to the kink of the in-work benefit; those with three children understood that their level of transfers was mostly determined by the number of children in their household, delayed the departure of their oldest children and increased their labour force participation such that they don't lose too much. Our findings also indicate that except single parents with one child at baseline, treated compliers are more likely to be the sole earner than untreated compliers, including when they re-partner.

This research contributes to different strands of literature that we largely discuss in Section II. Our paper was mostly inspired by the few examples in the literature on distributional effects of activation policies. In particular, our work is closest to the evaluation of a similar programme in the US by Alberto Abadie, Angrist, and Imbens (2002), as Matias D. Cattaneo, Jansson, and Ma (2021) also used these data to showcase their new method. In this paper, the authors use the κ weightings of A. Abadie (2003) to develop instrumental quantile treatment effects and showed large heterogeneous effects by gender. For women, Effects generally positive and higher at the bottom of the income distribution while men have no effect but at the upper end of the income distribution. D. H. Autor, Houseman, and Kerr (2017) also use this framework on a work-first job placement programme in the US. However, by focusing on the distribution of potential income and not quantile treatment effects, our approach does not need the often implausible rank invariance assumption such models require (Chernozhukov and Hansen 2008; Melly and Wüthrich 2017; M. Huber and Wüthrich 2019). Our methodological choices also reflect the recent working paper of Garbinti et al. (2023) using variations in bunching mass over time following a wealth tax reform under the Hollande presidency. To the best of our knowledge, this paper is the only other using instrumental variables to estimate bunching masses and elasticities and avoid parametric assumptions on the counterfactual densities.

Our results also relates to the analysis of the effect of the EITC on the poverty of single parents by Raj Chetty, Friedman, and Saez (2013). This paper uses differences in knowledge of the EITC which they approximate by the degree of bunching across counties. Using that in an event-study framework around birth of a child triggering eligibility, they show that the EITC reduces poverty. In another paper, Raj Chetty and Saez (2013) analysed the effect of providing information on the EITC through professional tax filers and also showed that complying tax-filers induced large bunching. However, we should emphasise the key difference in counterfactual. In the US, not knowing about the EITC means forgoing significant transfers. In our setting, not knowing about incentives is taxed, making the *make work pay* narrative fundamentaly wrong for single parents in France. Conversely, the information experiment on the Swedish EITC of Nyman, Aggeborn, and Ahlskog (2023) shows little to no reaction in a setting where the tax credit is automatic and separated for spouses.

If we attribute the primary effect of the programme to an enhanced understanding of the tax-benefit system, our findings underscore the significant tax burden faced by impoverished single-parent households, which serves as a potent disincentive at the extensive margin for some, while prompting a strong economic incentive for part-time employment in general. In either scenario, these incentives perpetuate situations where households lack sufficient income to escape poverty, thereby leading to a reliance on high levels of social transfers. The data suggests that among untreated compliers, increased work participation is not hindered by a lack of ability or opportunity, but rather by the burden of taxation.

Our work emphasises important contradictions in the tax-benefit system and politicians' narratives. For single parents, social transfers are both too low to lower poverty and deplete too fast to *make work pay*. Yet most do not know that. The government has very little incentives to make rules clearer for them, as reactions would most likely induce lower participation and working hours of single mothers, further increasing in-work poverty, gender inequalities in the labour market, and poverty among children. At the other side of the income distribution, Garbinti et al. (2023) analysed the effect of a wealth tax reform in France and also find strong reactions - lower wealth for more informed taxpayers. However, these results show that they most come from actual tax evasion and mis-reporting made simpler by the new tax regime. Conversely, low income families are more heavily controlled through algorithmic targeting, controllers are granted access to recipients' bank accounts and now, the administration receives direct feeds from payroll taxes, making income manipulations almost impossible and highly deterring. It also induces a large loss of freedom and privacy, as well as higher implicit tax rates (Quadrature du net 2023; Défenseur des droits 2017). Facing harder employment conditions and poor quality jobs (Rodrik and Stantcheva 2021), single parents on welfare really seem trapped in a poverty spiral from which it is hard to escape without drastic change in public policies.

Ultimately, there is a fundamental question on preferences for redistribution, social justice and optimal taxation (Blundell et al. 2009; Maniquet and Neumann 2021; Saez and Stantcheva 2016; Stantcheva 2021). Recent work show that social preferences are heavily politically polarised but do change with better knowledge (Kuziemko et al. 2015; Alesina, Stantcheva, and Teso 2018; Hvidberg, Kreiner, and Stantcheva 2023). A key driver of social preferences revolves around inequalities of opportunities and our work demonstrates that the tax-benefit system creates and exaggerates some of them. We hope that calling-out these injustices may help foster reforms that finally allows single parents to gain agency, means and freedom to improve their lives and that of their children.

The remaining of the paper is structure as Followed: Section II reviews the literature on welfare reforms and single parents, including our theoretical framework and a discussion on gender and household economics. The French tax-benefit system and the intervention are in Section III followed by the description of the data and a first set of descriptive results in Section IV. We discuss our identification strategy in Section V, present our main results in Section VI, and additional estimates in Section VII. Section VIII summarises and discusses our findings.

II Single parents and welfare reforms

In this section, we conduct a critical review of the literature in three parts. We begin by examining the economics of welfare reforms and recent academic debates surrounding the effects of the Earned Income Tax Credit (EITC). We explore the confounding effects of additional activation measures and the importance of factors such as stigma, information frictions, and costs. In section 2, we incorporate these constraints by adapting the typical model of labour supply with non-linear budget constraints. We then argue that current models and policies often overlook the impact of gender norms, which leads to an incomplete picture for a population like single mothers. Therefore, in part 3, we focus on understanding the specific constraints and incentives that women, mothers, and single parents may face when making decisions about labour market participation.

1 Welfare reforms and the labour market

The literature on this topic originated in the 1960s and early 1970s, when the static model of labour supply was applied to the work incentives of a negative income tax[4]. In their review, Chan and Moffitt (2018) underline how the international literature has been closely tied to policy developments and to welfare reforms in different countries. Many economic theories were built on reforms and new reforms were inspired by the diffusion of economic ideas[5]. While models usually build on genderless agents, It is worth noting that a significant share of the literature was concerned about the effects of welfare on fertility and single-parenthood (Robert A. Moffitt 1998 ; Ellwood 2000).

Typical welfare reforms with an activation orientation introduce in-work benefits[6], increase them up to a point in the wage distribution, add work requirements, sometimes in systems with variations in the marginal tax rates. Figure II.1 reproduces the illustrations of such reforms from Chan and Moffitt (2018). In this typology, the 2019 reform of in-work benefit in France corresponds to the Panel b, moving-up monetary incentives with a necessary flatter slope after the kink C′. These four panels conveniently represent the typical behavioural reactions that such welfare reforms potentially introduce. The point being, with heterogeneous agents, reactions can go in many directions and the overall effects can hide sizeable reactions at the intensive margin. The first generation of static models were already very clear about that (See e.g. Burtless and Hausman 1978; Robins 1985). Theoretical models predict that the extensive margin decision depends on taxes and transfers at the desired level of earnings, *i.e.* the intensive and extensive margin decisions are inter-dependent (Eissa, Kleven, and Kreiner 2006).

The effect of the EITC: A controversy The Earned income tax credit plays a special role in the academic literature and political discourses on *making work pay*, inspiring many similar schemes in the world - including France (See Section 1). A large literature documents a sharp rise in the employment of single mothers, especially single mothers with two or more children, following the 1993 EITC expansion for these family types, and concludes that reactions at the extensive margin strongly dominate those at the intensive margin. (Meyer and Rosenbaum 2001; Grogger 2003; Eissa and Hoynes 2006; Herbst 2011; Hoynes, Miller, and Simon 2015).

However, there is an academic debate on the validity of these estimates[7]. In a first working paper, H. Kleven (2019) reconsiders the data and the many reforms of the EITC. He estimates hundreds of difference-in-differences considering different reform episodes, different samples, different comparison groups, different extensive margin measures, and different control variables. Allowing for all possible permutations of specification choices, the estimates are symmetrically distributed around zero, except those on the 1993 reform. For the latter, the distribution of estimates is shifted to the right and has a mean elasticity of 0.63. Leaving aside the 1993 reform, the main message from this paper is that the EITC has no effect on labour *per se*. A conclusion challenged by Whitmore Schanzenbach

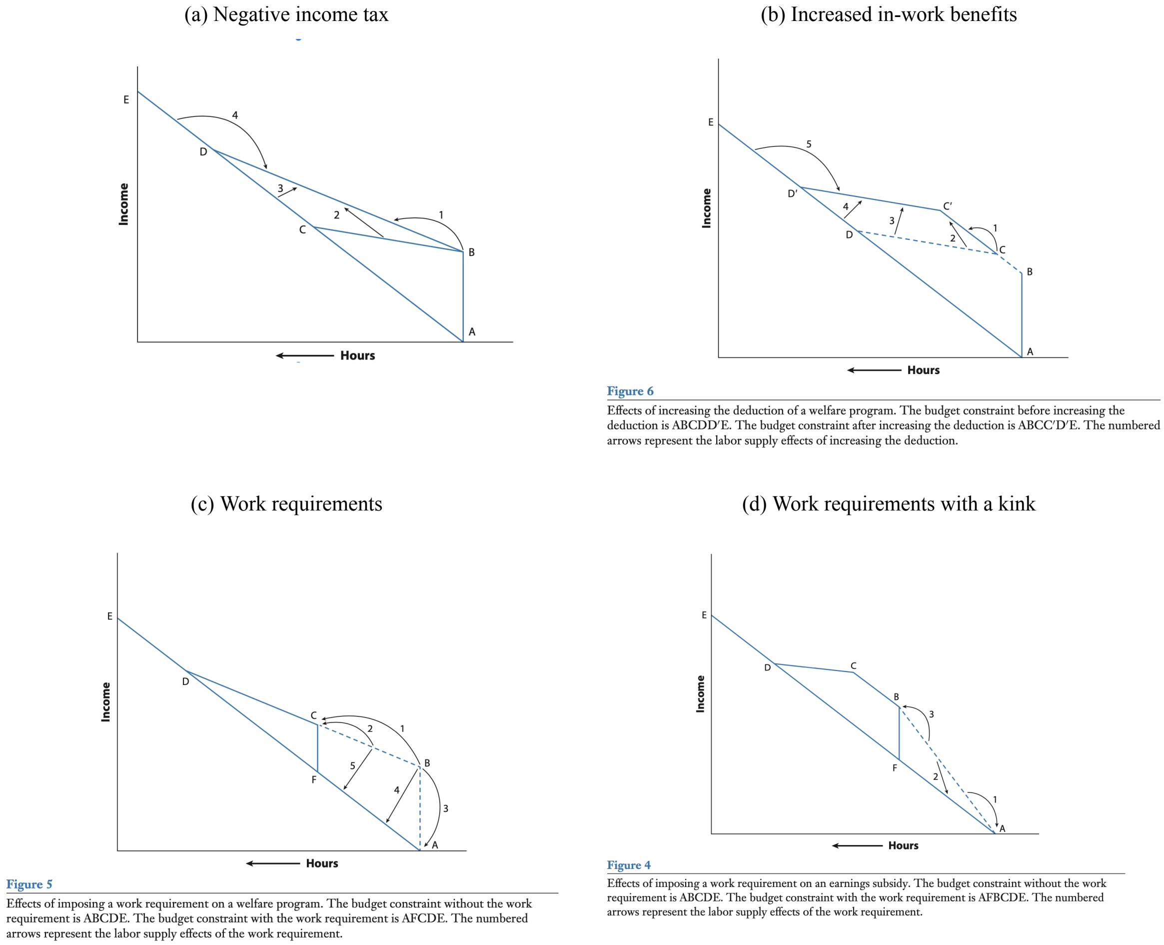
[4] Burtless and Hausman (1978); Diamond (1998). See R. Moffitt (1992) for a review of this early literature.

[5] For instance, many reforms were directly influenced by the negative-income tax model of Friedman, and many new models came after some new policies were implemented. See Robert A. Moffitt (2003) for a specific discussion of the role of Friedman's model in shaping welfare reforms around the world.

[6] In-work benefits can be tax credits or cash transfers.

[7] Identification and methodological challenges are also important. In particular, they mostly use difference-in-differences where treatment status is based on fertility, a very strong assumption considering the impact of children on labour market outcomes *per se* (H. Kleven, Landais, and Leite-Mariante 2024). Regressions aggregating different groups treated at different times may be biased by negative weightings in two-way fixed-effects regressions (Goodman-Bacon 2021).

Figure II.1: A typology of welfare reforms from Chan and Moffitt (2018)



and Strain (2021), who propose a similar re-analysis of the many waves of extension of the EITC, with different samples and including controls for the business cycle, a larger observation window and scope of the analysis with non-federal schemes. While Kleven uses labour participation the previous week (the official ILO measure), they choose labour market participation over the year, arguing that the EITC depends on yearly labour income. They also argue that the three-way fixed effect interaction between state x time x children of Kleven absorbs the effect. In a revised version, Henrik Jacobsen Kleven (2023) argue that the specifications of Whitmore Schanzenbach and Strain (2021) are "*strong outliers in the distribution of estimates across a wide range of specifications*".

There are four main arguments explaining why estimates for the 1993 expansion of the EITC are different from others. First, it is strongly confounded by welfare reforms occurring from 1992 to 1996, including the Temporary Assistance for Needy Family (TANF) programme and the many welfare-to-work programmes of the 1990's. Second, it was bigger and more advertised, which matters in a world with optimisation friction. Third, the economy was booming, which may have had heterogeneous effects on single women with and without children, especially at a time where welfare reform were putting strong pressure on single mothers to find employment. Fourth, significant changes regarding social norms towards welfare assistance and work were documented in the 1990s (Peterson 1997; Ellwood 2000; Robert A. Moffitt 2015).

Welfare-to-work programmes and other activation measures Monetary incentives are generally conditioned upon participation in other active labour market activities, including welfare-to-work. The economic motivations for activation concern 1) human capital, 2) frictions in the labour market, and 3) the moral hazard effects of unemployment insurance and welfare provision (Crépon and van den Berg 2016a). Researchers and policymakers have implemented various *carrots vs. sticks* schemes to foster participation in ALMP, insisting on a "threat" or "pull" effect, with varying results (Arni, Lalive, and Van Ours 2013; Hohmeyer and Wolff 2018; Avram, Brewer, and Salvatori 2018; Morescalchi and Paruolo 2020). They have been widely implemented with many variations in policy instruments, target groups and context. As a result, the econometrics literature has dramatically improved in quality (J. J. Heckman, Lalonde, and Smith 1999; Rothstein and von Wachter 2017, Mars) and has been extensively reviewed in several systematic reviews (Card, Kluve, and Weber 2010 ; Filges et al. 2015; Card, Kluve, and Weber 2018; Vooren et al. 2019b). They generally exhibit positive effects, but with considerable variations in effect sizes. Results between systematic reviews also vary depending on inclusion criteria, especially regarding the quality of the research design and target group.

However, the literature on the effects of welfare-to-work on single parents is very contrasted, and usually more negative (Smedslund 2006; Gorey 2009; Campbell et al. 2016b; Gibson et al. 2018). There is a strong discrepancy between the numerous experimental research in North America and their scarcity in Europe. As a consequence, most European evaluations rely on stronger hypotheses and may be far less reliable (Heim 2024). For instance, the review from Bergemann and Van Den Berg (2008) on ALMPs for women in Europe concludes that they are very effective but only 4 out of 39 included research use a randomised experiment, two of which show negative results. There is also a large heterogeneity on the type of intervention and treatment effects. For instance, Alberto Abadie, Angrist, and Imbens (2002) use instrumental quantile regression on a randomised welfare-to-work programme in the US and find strong positive effect concentrated at the bottom of the distribution for women. Conversely, Mogstad and Pronzato (2012) analyse the Welfare reform in Norway which introduced in-work benefits and workfare requirements and show average positive effects on single mothers on the one hand, and large negative effects for those at the bottom of the distribution. This heterogeneity may also hide exacerbated psychological costs in that the mandatory aspect of programmes is frequently experienced by participants as a loss of autonomy, especially when said programmes are deemed inadequate for those they're supposed to help. The systematic review of Shahidi et al. (2019) concludes that social assistance programmes in high-income countries are insufficient in preserving the health of socio-economically disadvantaged populations, indicating that the scope and generosity of existing programmes fall short in compensating for the negative health consequences associated with poverty.

In France, around 1/4 of all minimum aid recipients (of which 54% are women) cite health issues as their main obstacle to employment (Cour des comptes 2022). Policymakers therefore face an asymmetric information problem: they do not know which part of the population cannot work and which part can. The definition of *being able to work* is a social construct and active labour market policies typically changed these definitions[8]. When imposing workfare obligations, policies can "*miss*" their targets in two ways: i) by futilely mobilising or penalising people who cannot work; ii) by forcing people who do not need them to take part in these programmes. The efficacy of mandatory welfare-to-work programme therefore depends on the size of each group and raises important ethical questions. For Molander and Torsvik (2015), mandatory workfare cannot be justified if such groups exist in the eligible population, and should only exist when they target specific populations.

These considerations were formalised by Kreiner and Tranaes (2005) in a model incorporating people who are in work, unemployed or not looking for work, and where it is not possible (or costly) to observe job-seeking efforts. One consequence of this configuration is that the risk of involuntary unemployment is under-insured. By incorporating obligations and constraints - even unproductive ones - the model predicts better coverage of the risk of unemployment for workers. Mandatory welfare-to-work then acts only as a screening mechanism where the constraint reveals the latent types of individuals. If these policies are sufficiently effective to produce the expected incentives, then they allow better targeting of minimum social benefits' recipients. Pavoni and Violante (2007) propose an *optimal welfare-to-work* model considering together unemployment insurance, workfare and welfare. The optimal design they propose includes decreasing replacement rate of the unemployment insurance, in addition to job-search monitoring and support, training and career orientation plans upon exhaustion, and a basic income schemes for those who *cannot* work. The latter point is crucial in acknowledging that some individuals may be unable to hold a job and should not be imposed costly activation measures.

[8] Imposing workfare obligation to claimants of the disability allowance for instance. See for instance the evaluation of the programme *pathway to work* in the UK from Adam et al. (2008).

Non take-up of social programme At this point, it is useful to add that in-work benefits and tax credits are not directly provided and require people to submit their incomes and accept the conditions under which such benefits are offered. The application processes may be highly taxing *per se*. Welfare reforms and activation are strongly linked to a logic of reducing and rationalising public spending, driven both by public debt concerns and the rise of the New Public Management paradigm, which first gained prominence in English-speaking and Scandinavian countries (White 2019). This led to a strengthening and institutionalisation of the logic of evaluating policy implementation and outcomes (Lacouette Fougère and Lascoumes 2013; Bozio 2014; Fougère and Heim 2019; P.-H. Bono et al. 2021), but also to the multi-faceted complexification of administrative processes. This directly connects to the notion of *"administrative burden"*, and how seemingly parametric measures may have deep economic, political and social impacts. Administrative burden are composed of three components: Learning costs, compliance costs and psychological costs[9]. According to Herd et al. (2023) in the US case, burdens are "*policymaking by other means*" that have have large effects on access to rights and public services, facilitate social control and reinforce inequality. Moreover, their effects accumulate over time and people with fewer resources are less equipped to manage burdens.

For policymakers, more rigorous screening process may improve targeting efficiency, but the associated complexity is costly to applicants. Henrik Jacobsen Kleven and Kopczuk (2011) rationalise this in a model characterising optimal programmes when policy makers choose benefit level, eligibility and screening intensity. Consistent with many real-world programmes, optimal choice for policymakers feature high complexity, incomplete take-up, classification errors of both over-rejection and excess-award.

In the field, reducing low take-up has proven to be very hard to achieve, generally leaving the most vulnerable behind. Bhargava and Manoli (2015) analyse the role of "psychological frictions" in the incomplete take-up of EITC in a field experiment on 35,050 tax filers who failed to claim $26 million despite an initial notice. The information of the unclaimed benefit led to substantial additional claiming but attempts to reduce perceived costs of stigma, application, and audits did not. Linos et al. (2020) ran six pre-registered, large-scale field experiments to test whether "nudges" could increase EITC take-up (N=1 million). Despite varying the content, design, messenger, and mode of their messages, they find no evidence that they affected households' likelihood of filing a tax return or claiming the credit. Similarly, Nyman, Aggeborn, and Ahlskog (2023) ran an information experiment randomising 37 000 leaflets providing information on the Swedish EITC in a pre-registered randomised experiment and also find precise null effect both at the extensive and intensive margin. Similar results were found in France by Chareyron, Gray, and L'Horty (2018) in a mail experiment on 4000 new welfare claimants to foster registration at the employment agency for welfare recipients. The variations in the mails have no effect on registration. Finkelstein and Notowidigdo (2019) run a field experiment over 30 000 elderly likely eligible for the Supplemental Nutrition Assistance Program (SNAP), either provided with information that they are likely eligible, provided with this information and offered assistance in applying, or are in a "status quo" control group. 6% of the control group enrolled over the next 9 months, 11% in the information group and 18% in the information with assistance. Evidence suggest optimisation frictions greater for needier individuals as compliers tend to have higher net income and are less sick than the average enrollee. In France, Castell et al. (2022) ran two field experiments to foster access to all social benefits. In the first experiment, they show that receiving assistance increase take-up of new benefits by 30% while the second only use an online simulator through an information experiment. Using a marginal treatment effect analysis, they show that benefits are mostly concentrated among the least likely to attend and suggest that transaction costs deter eligible people from applying to benefits and from accessing government's assistance to help them apply.

However, both Finkelstein and Notowidigdo (2019) and Castell et al. (2022) show that human assistance and personalised information do reduce non-take-up of social benefits. Similarly, Raj Chetty and Saez (2013) also ran an experiment in which tax pre-payers provided simple but personified information on the EITC to half of their clients, reaching a total sample size of 43000 households. They show that about half of the tax-prepayers induced their clients to chose earning closer to the kink and 10% less likely to have very low income compared with control. In a two phase randomised control trial of the *moving to opportunity* programme in the US, Bergman et al. (2019) show that providing high-intensity, customised support *i.e.* clear and tailored information about high-opportunity areas, short-term financial assistance, customised assistance during the housing search process, and connections to landlords - increased the moving rate from 15% in the control group to 53% in the treated group. The second phase

[9] "Learning costs arise from engaging in search processes to collect information about public services, and assessing how they are relevant to the individual. Psychological costs include the stigma of applying for or participating in a programme with negative perceptions, a sense of loss of power or autonomy in interactions with the State, or the stresses of dealing with administrative processes. Compliance costs are the burdens of following administrative rules and requirements." (Moynihan, Herd, and Harvey 2015)

of the experiment tests each component of the high intensity support to disentangle which mattered most. Results show that information yields treatment effects 5 times lower than the combined intensive treatment and reduced support with information 3 times lower. It is the combination of high intensity support with clear and tailored information that works.

Welfare stigma and compliance costs In the early literature, R. Moffitt (1983) introduced welfare stigma as a cost in the utility function, generating non-take-up at equilibirum. More recently, Kline and Tartari (2016) built a model with both welfare stigma and compliance cost affecting utility and participation in the programme. This model incorporates the "hassle" associated with welfare work requirements and also allows agents to manipulate their reported incomes at a cost. Using this model, they are able to make restrictions on preferences to offer informative bounds on the labour supply response of single mothers in a randomised experiment of a welfare-to-work programme in the US. This programme strengthened work requirements and increased sanctions for welfare recipients who fail to seek work. Moreover, it changed the manner in which welfare benefits phase out by disregarding earnings up to an eligibility threshold (or "notch") above which benefits abruptly drop to zero. These incentives push at least 20% of control group women whose earnings are above the notch to reduce their earnings below the notch (but remain working) and receive welfare under the experiment.

The stigma associated with social benefits is very pervasive. In a behavioural economics experiment manipulating the social image of access to a benefit, Friedrichsen, König, and Schmacker (2018) show that the stigma associated with both "living off others" and receiving aid "for the least qualified" causally reduces the intention to use this aid. Celhay, Meyer, and Mittag (2022) compare participation in social programmes reported in declarative surveys with information on participants' networks and document strong under-reporting and manage to exclude mechanisms other than stigma to explain it.

In France, while we can observe a steady increase in the number of welfare recipients since the 2008 reform[10] (Hannafi et al. 2022), estimates of non-take-up range from 28% to 35% of eligible people (Domingo and Pucci 2013). In a qualitative analysis of non-take-up, Chareyron (2018) mentions that 14% of eligible non-recipients cite not knowing about the aid as the main reason why they don't resort to it, and 71% that they do not know how the amount is calculated. To those 85% ascribable to learning costs are added 23% associated to stigma, or psychological costs, although research has yet to identify exactly how dissuasive they are. Interestingly, since having a child opens rights to benefits and therefore facilitates contact with the relevant institutions, single parents are proportionately more numerous within beneficiaries (7%) than non-takers (under 3%). On the control of welfare recipients, practices of the administration in charge (*CNAF*) towards single parents – mostly women – have been subject to heated controversies in the past years: in his analysis of the institutional evolutions of social benefits, Dubois (2021) presents the institutionalisation and hardening of control policies and practices since the 1990s as part of a "symbolical and moral economy" where "assistantship" (*assistanat*) and fraud are the necessary deviances to the promotion of the norm, i.e. *la valeur travail*. Dubois' conclusions that these evolutions mostly penalise vulnerable populations, in particular the disabled, migrants and single mothers, were further supported by new evidence regarding CNAF's practices in scoring recipients' likelihood of fraud (Quadrature du net 2023). Like many administration around the world, they use predictive models to flag files with higher risk of *errors or fraud*, without distinction. The model is a simple Logit where variable have been previously selected using Lasso. The thorough analysis of the source code by Quadrature du net (2023) shows that the most impoverished beneficiaries consistently have a higher suspicion score. Elements such as separations or single-income households increase the score in a way that it is hard not to conclude to the targetting of single parents, leading many to argue that CNAF's control policy disproportionately sanctions vulnerable people going through a particularly tough time. Beneficiaries are also required to notify any change in relationship status, and not doing so is akin to fraud. In spite of criticisms from for instance the *Défenseur des droits* (Défenseur des droits 2017), the concept of "marital situation" is vague enough that non-financially enmeshed people can still see their rights revoked. For single mothers, such intrusiveness considerably (needlessly?) raise the compliance and psychological costs.

[10] See Subsection 1 for a brief description of the French welfare system and Appendix A for the history of welfare reforms in France and a short review of the literature on the effects of monetary incentives in France.

2 Tax salience, information and optimisation frictions

In this section, we lay down the utility maximisation model used in public economic literature on bunching that rationalises intensive margin reactions around variations of tax rates[11]. As our presentation goes, we discuss how the literature introduces the frictions we discussed earlier.

The bunching model Consider individuals indexed by i, whose heterogeneity can be summarised by scalar variables a_i and ψ_i, representing abilities and adjustment costs, respectively. Typically, ψ_i can be interpreted as monetised stigma or costs associated with changing situations (Kline and Tartari 2016). Later, we will explore how these parameters might be endogenous to the tax-benefit system or active labour market policies. For now, we abstract from gender and family structure, following a simplified model of optimisation friction.

Individuals maximise their utility U by choosing disposable income Y through optimising pre-tax earnings X, subject to a budget constraint that includes mean-tested social transfers, leading to kinks in the implicit tax rate.

We consider the typical iso-elastic utility function:

$$U(Y,X) = Y - \frac{a_i}{(1+\frac{1}{e})}\left(\frac{X}{a_i}\right)^{1+\frac{1}{\varepsilon}} - \psi_i \tag{3.1}$$

where $Y = X - T(X)$ is the net disposable income and ε is the elasticity parameter, constant across individuals.

This model disregards factors like children or spouses, which do not pass a basic reality check. Both are endogenous, affect tax rates, and spouse incomes are part of the household income Y but may not enter a parent's utility function as such. We come back to this point in subsection 3. Despite these important simplifications, this is the structural model commonly used in estimating labour elasticity of single parents, particularly around kink points of programmes like the EITC (Saez 2010; Raj Chetty, Friedman, and Saez 2013).

To formalise the optimisation problem, we express the budget constraint in linear form: $X = (1-\tau')X + R$, where $\tau' \equiv T'(X)$ is the (perceived) marginal tax rate and $R \equiv T'(X)\cdot X - T(X)$ is virtual income[12]. In a tax-benefit system without kinks, the first-order condition yields the optimal pre-tax income[13]:

$$X_i^* = a_i(1-\tau)^{\varepsilon}$$

Taking this expression in logs, we get $log(X_i^*) = \varepsilon \log(1-\tau) + \log(a_i)$, which makes clear that the parameter ε measures the elasticity of pre-tax incomes to the net-of-tax rate.

In the presence of a kink K with marginal tax rates τ_a below X^* and τ_b above X^*, workers who would chose a level of pre-tax income above the kink in its absence now chose $X = K$ because they face a higher marginal tax rate. Among the bunchers, the one with the highest ability a_i is the one who would chose exactly $X = K$ if the higher tax rate τ_b prevailed below the kink point. This individual is the *marginal buncher* and his location is given by

$$\Delta X^* = K(1-\tau_b)^{\varepsilon}$$

All individuals with ability between $\frac{K}{(1-\tau_a)^{\varepsilon}}$ and $\frac{K}{(1-\tau_b)^{\varepsilon}}$ bunch at K.

The resulting distribution has a mass-point at K but is otherwise continuous, as presented in Figure II.2. An important result from Saez (2010) is that for small tax rate changes, we can relate the elasticity to the earnings change ΔX^* for the individual with the highest *ex ante* earnings who bunches *ex post*:

$$\varepsilon = \frac{\Delta x^*/x^*}{\Delta\tau/(1-\tau_a)}, \tag{3.2}$$

[11] See for instance Blundell et al. (2009);Saez, Slemrod, and Giertz (2012);Bierbrauer, Boyer, and Peichl (2021);Henrik Jacobsen Kleven (2016);Kostøl and Myhre (2021)

[12] The intercept of a linear budget set passing through the point $(x, T(X))$. See Piketty and Saez (2013).

[13] To see why, note that the marginal benefit of a variation in earnings is given by 1-τ', while the marginal cost is $(\frac{x}{a_i})^{\frac{1}{e}}$). At equilibrium, both equate and yield the given solution.

where $\Delta\tau = \tau_b - \tau_a$ and ε is the elasticity of pre-tax income to variation in the marginal tax rate. The higher the elasticity and the change in taxes at the kink, the larger is the range ΔX^* of bunchers.

Figure II.2: Bunching at kink point (Kleven 2016)

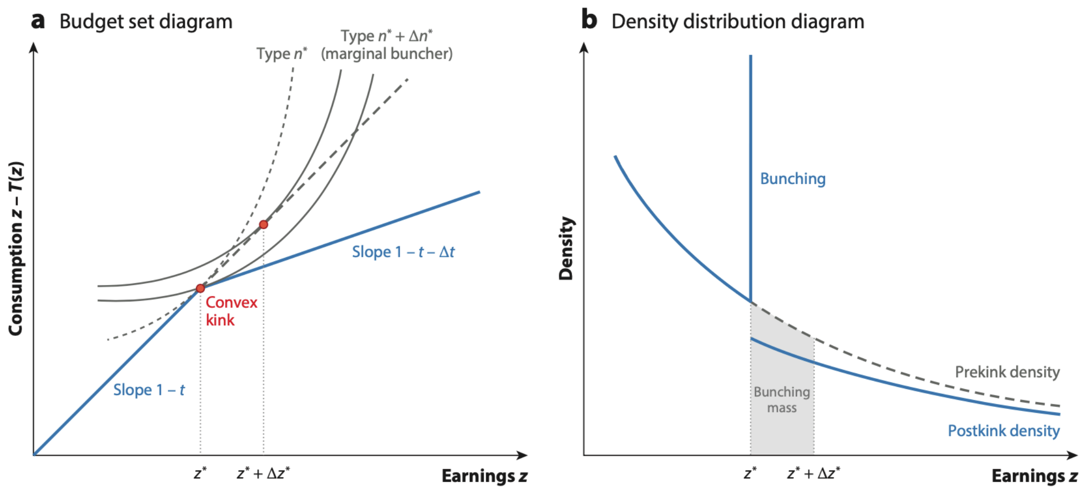


Figure 1

Kink analysis, showing the effects of a convex kink—a discrete increase in the marginal tax rate from t to $t + \Delta t$ at the earnings threshold z^*—in a (*a*) budget set diagram and (*b*) density diagram. In panel *a*, the individual with ability $n^* + \Delta n^*$ is the marginal bunching individual. This individual chooses $z^* + \Delta z^*$ before the kink is introduced and z^* after the kink is introduced. All workers initially located on the interval $(z^*, z^* + \Delta z^*)$ bunch at the kink, whereas all those initially located above $z^* + \Delta z^*$ reduce earnings within the interior of the upper bracket. As shown in panel *b*, the implications of these responses for the earnings distribution are sharp bunching at z^* (the size of which is equal to the *gray shaded area* just above z^*) and a left shift of the distribution in the upper bracket.

Identification with bunching design The general idea behind this model is that the size of the bunching mass may be informative of the underlying elasticity, but requires several additional information to be identified. The estimation methods proposed by R. Chetty et al. (2011) or Raj Chetty and Saez (2013) rely on an affine approximation of the counterfactual density in a chosen interval defining the bunching region[14].

The bunching mass is defined by:

$$B^* = \int_{X^*}^{X^*+\Delta X^*} f_{Y0}(u)du \approx f_{Y0}(K) \cdot K \left(\left(\frac{1-\tau_a}{1-\tau_b} \right)^{\varepsilon} - 1 \right)$$

where the approximation is that proposed by R. Chetty et al. (2011).

The problem is that this estimator is sensitive to the width of the bunching window and the quality of the parametric fit outside the bunching region. With kinks, the interval length is unknown and could be large because it depend on the elasticity we want to estimate. In Fact, Blomquist and Newey (2017) show that there is no way of inferring the elasticity parameter from bunching mass and continuity alone. This is particularly true for kinks where there is no way to know the position of the marginal buncher without assumptions on the other parameters.

The polynomial fit typically uses data outside both side of the kink to fit the implicit ability distribution a_i. In a few cases, researchers use a control group to infer the missing distribution, but not in the context of welfare policies[15]. Another approach consists in modelling or manipulating information to generate reactions by reducing frictions.

[14] Saez (2010) uses a trapezoidal approximation and implicitly assumes an affine function over the bunching interval. See Bertanha et al. (2023) for a very pedagogical discussion.

[15] See for instance Coles et al. (2022) on corporate tax, Gelber, Jones, and Sacks (2020) on age retirement and Garbinti et al. (2023) on wealth tax.

Optimisation frictions So far, the model assumes no friction and immediate reaction to the tax rate. A first type of optimisation friction arises for individuals around kinks and notches for whom adjusting to incentives may be too costly. In such case, the bunching mass is diffuse and identification of the bunching region is even more sensitive. In the presence of optimisation frictions, the observed elasticity does not correspond to the structural elasticity and must be modelled explicitly.

For instance, Kostøl and Myhre (2021) introduce a "*knowledge*" parameter in the household budget constraint such that part of the population does not know about the change in tax rate at the kink and optimises as if the tax rate was linear at the kink. Formally,

$$T(x) = \tau_a(X_i) + \theta_i \tau_b (X_i - X) \cdot \mathbb{1}(X > X)$$

The parameter θ_i measures the degree to which agent i wrongly perceives the tax rate for incomes over k. Like before, the utility maximisation yields the optimal pre-tax income with frictions X_{Fi} which now takes the following form:

$$X_{F,i} = \begin{cases} a_i (1 - \tau_a - \theta_i \tau_b \cdot \mathbb{1}(X > k))^{\varepsilon} & \text{if } U_i(Y, X_F) - U_i(Y, X^*) \geq \psi_i \\ a_i (1 - \tau_a)^{\varepsilon} & \text{if } U_i(K, X_F) - U_i(K, X^*) < \psi_i. \end{cases} \tag{3.3}$$

This equation shows that both adjustment costs ψ_i and information frictions θ_i reduce the impact of monetary incentives. Only agents with costs lower than the benefits of adjusting, and individuals who are aware of the kink, respond. It also emphasises the interdependence between these two types of frictions. Increasing the proportion of individuals with the correct understanding of the tax also changes the perceived benefits of re-optimisation. We come back to this when we present our empirical strategy in section 2.

Kostøl and Myhre (2021) use this model to analyse a tax reform in Norway together with an information experiment providing details on the reform and the existence of notches. Consistent with this model, their results show large adjustments at the intensive margin, twice higher in the group that received information than in the control group. Raj Chetty and Saez (2013) use a similar approach in the information experiment discussed earlier and also find much larger elasticities in the group who received information and support. Caldwell, Nelson, and Waldinger (2023) analyse the effect of uncertainty around the amount of EITC and show that they distort individuals' consumption-savings choices enough to cause welfare losses among EITC filers on the order of 10 percent of the value of the EITC. In Denmark, welfare recipients are required to work 225h in the past 12 months and face a reduction in their monthly payments if they fail to comply with the work requirement. In a field experiment, Cairo and Mahlstedt (2023) randomised i) access to a personalised online tool that offers continually updated personalised information including the number of working hours they have accumulated and their personal deadline for compliance with the work requirement; ii) notification messages that are almost identical to those available in personalised tool, but workers assigned to the message treatment do not gain access to the online tool iii) a control group neither receiving notification messages nor have access to the online tool. Their results show heterogeneous reaction depending on individual initial situation and treatment received. Personalised information increases employment for those who initially fail to meet a work requirement at the start of the intervention while notifications decrease the labour supply of workers with limited incentives to work extra hours in the immediate future.

Raj Chetty, Friedman, and Saez (2013) have a rather unique approach in this literature combining a natural experiment with bunching to recover the effects of the earned income tax credit on labour market participation of single parents and poverty. Like Saez (2010), they observe sharp bunching at the first kink of the earned income tax credit, but with significant variation across counties, which they interpret as a lack of *knowledge* of the EITC. They use an interesting identification strategy, essentially using bunching as a first stage of the effect of EITC on the distribution of incomes of single mothers around birth. Assuming that differences in bunching by self-employed individuals around the first kink of the EITC around child birth across counties only reflect knowledge of EITC, they leverage differences in i) the share of bunchers ii) by relative months from birth iii) between high and low knowledge county, where the shifting share of bunchers after birth in high knowledge counties rescales the change in disposable incomes around birth. Their findings suggest substantial intensive margin earnings responses to EITC incentives, conditional on knowledge and showed that the EITC increased incomes at the bottom of the distribution.

Theses results have important implications. The very existence of this natural experiment shows that salience matters and that a large share of the population does not know about programmes intended for them. Interestingly, while the informational and psychological frictions associated with non-take up of social programmes are widely acknowledged in the literature (van Oorschot 1991; Portela et al. 2022; Hannafi et al. 2022; Ko and Moffitt 2022), they have been used mostly as an explanation for not observing any intensive margin responses. However, theoretical models of intensive and extensive margin responses (see e.g. Eissa, Kleven, and Kreiner 2006) imply that such frictions are equally important for the extensive margin.

The elasticity of women labour supply is an important question for which there does not seem to be a definitive answer despite important sophistications in the model used. For instance, O. Attanasio et al. (2018) shows that there is substantial heterogeneity in women's labour supply elasticities at the micro level and build a life-cycle model to aggregate elasticities. In France, Briard (2020) reviews the literature and shows that most estimates use structural models with discrete choices but lack consistency with regards to their spouse's income.

It is also worth noting the choice of vocabulary used in models of labour participation of single parents on welfare. Most of the work on the EITC consider single parents and all model their decision as "*tax evasion*". Agents maximise consumption, leisure time and minimise tax. The models trying to account for stigma models single mothers as *genderless tax-avoiders*, which is not without irony. This brings us to what we think is the most important critic of this literature: the lack of consideration for the specifics of the targeted populations and the context of decisions.

3 Welfare reforms and the economics of gender, family and human capital

***Gender trouble* in economics** Feminist economics criticised the rationality framework used in mainstream economics focusing on the individual characteristics of lone mothers, which is, at best, conceptualised as human capital (education, training, and experience), individual resources (e.g. income), and constraints (e.g. number and age of children). However, what is economically rational is different in different welfare state regimes which in turn determines their origins in collective political and ideological views of society, particularly concerning the relation between individuals, families, states, and markets (S. Duncan and Edwards 1997; Pollmann-Schult 2018). The source of economic rationality, therefore, at least in this case, partially lies outside the market and in the domain of collective and highly gendered understandings about proper social behaviour and the internalisation of such norms in personal values (Stavrova and Fetchenhauer 2015; Hakovirta, Kallio, and Salin 2021; Gong, Stinebrickner, and Stinebrickner 2022; Reich-Stiebert, Froehlich, and Voltmer 2023). For instance, in England, Cavapozzi, Francesconi, and Nicoletti (2021) find that having peers who embrace egalitarian gender norms leads mothers to be more likely to engage in paid employment and have a greater share of total paid hours within their households. These effects are particularly pronounced among less educated women. They estimate that about half of the impact on labour force participation is due to women conforming to their peers' gender role attitudes, while the other half is attributed to the peer behaviour spillover effect on the labour market. A similar result was demonstrated in France by Maurin and Moschion (2009) using an instrumental variable strategy. Starting from the premise that the gender of the first two children affects the likelihood of having a third child and has a negative reduced-form effect on employment and income (J. Angrist and Evans 1998), Maurin and Moschion (2009) show that the presence of more or fewer families with two children of the same gender in the neighbourhood (and therefore, mothers working less due to this) reduces other women's labour force participation. When there is participation, it remains constrained by gendered household organisation : for instance in France, Le Barbanchon, Rathelot, and Roulet (2020) shows that when considering a job, women trade-off commuting distance and wage rates.

In order to understand such rational, economists must take into account the wide range of such trade-offs that mothers have to deal with, in spite of their generally implicit nature. Not least among them is the fact that time spent with children, particularly in early childhood, remains the least replaceable and one of the two most valuable investments in human capital that parents can provide, together with income (Elango et al. 2016; Doyle 2020; G. Duncan et al. 2023).

Moreover, and although definitive evidence is hard to obtain, many economists agree that maternal and paternal times are more complements than substitutes (as is the case, for instance, with the Cobb-Douglas form adopted by Del Boca, Flinn, and Wiswall (2014) or Goussé, Jacquemet, and Robin (2017)). Education, fertility, marriage, investment in children education and employment choices are increasingly analysed together ; and men and women play different roles in the models (Pierre-André Chiappori, Radchenko, and Salanié 2018; Goussé, Jacquemet, and Robin 2017). Bargaining in the household is not a negligible aspect. For instance, the analysis of the effects of unilateral divorce laws in the US showed that it did not increase divorce beyond a very short-lived bump (Wolfers 2006). However, Stevenson and Wolfers (2006) showed that they reduced female suicide by 8 –16 percent, domestic violence by roughly 30 percent for both men and women, and feminicides by 10 percent.

On the economics of households More sophisticated theories have been developed, incorporating, for example, strategic, cooperative, or non-cooperative interactions among household members, preferences, and transferable utility that may depend on (gendered) reference norms, which are more compatible with the data. Within mainstream economics, new perspectives on gender have thus emerged (Bertrand 2011). Yet, labour economics remained unaffected by these evolutions in its cousin field of household economics. In the bunching literature, when researchers define utility of the households, they usually use the unitary household with separable utility[16] although this model has been consistently rejected in the data (Lundberg and Pollak 1996; Pierre-Andre Chiappori and Mazzocco 2017).

In our presentation, we defined X as pre-tax income, living the possibility that the latter comes from single mothers's own income or that of a potential spouse, and depend on the family structure. If we introduce children and spouse as choice variables, the tax rates also vary across conditions and depend on joint spouse incomes. This would require a much more sophisticated model to represent the household's decision making. Such formalisation is out of the scope of this research but emphasises the complexity and endogeneity of many choices, all dynamic with lasting effects. Among them, the many choices and shocks that led them to be poor single parents.

Bertrand et al. (2021) develop a model with nested generations in which women decide whether or not to acquire skills and then, based on the partner they obtain, decide whether or not to marry and, in a non-cooperative manner, decide how to allocate their time between work, leisure, and the production of a "public good", *i.e.* children's education. The marriage gap between qualified and non-qualified women emerges endogenously due to higher wages resulting from their degrees and the different time allocation decisions these higher wages generate. The key force behind the emergence of this marriage gap lies in gender norms that create marital disagreements about the level of education to dedicate to children. The model predicts a negative relationship between the mother's education and the probability of marriage until the return on education is such that it becomes attractive for men to be with them or when gender norms are more egalitarian. The significant increase in women's education levels worldwide since the 1960s, along with access to technologies reducing the time required for domestic tasks or enabling fertility control and the emergence of care outsourcing services, drastically altered the balance within families (Claudia Goldin 2006-May). However, despite these major changes, gender specialisation in domestic or educational tasks has only partially evolved, while structurally modifying gender relations in society (Juhn and McCue 2017).

There is also a rather large literature showing that gender norms and frictions within households are sometimes the very cause of single parenthood to the point that researchers used manifestations of such frictions as instrument of single parenthood. For instance, Bedard and Deschênes (2005) use the gender of the first child (assigned female at birth) as an instrumental variable to predict a higher risk of separation[17]. They measure its effects and find that women separated because of having a girl as their first child have significantly higher average incomes than if they had not been separated. This effect is driven by increased labour force participation and higher working hours. Elizabeth O. Ananat and Michaels (2008) use the same instrument to analyse the effects of separation on income distribution quantiles. Their results indicate that separation increases the share of women in the lower and upper tails of the income distribution, meaning higher shares of impoverished and affluent women, resulting in negative income impacts. Simultaneously, divorce increases the share of women in the upper echelons of the distribution and boosts their income levels. Frictions on incomes are also a major cause of separation and many research find a sharp discontinuity in the density of couples where mothers earn more than their partners, also associated with a larger happiness gap and less leisure for women (Krueger 2008; Stevenson and Wolfers 2009; Bertrand, Kamenica, and Pan 2015; Bodson and Kuépié 2012; Lippmann, Georgieff, and Senik 2020; Flèche, Lepinteur, and Powdthavee 2020; Blanchflower and Clark 2021).

[16] See for instance the work of Nyman, Aggeborn, and Ahlskog (2023) on information on the Swedish EITC

[17] see also Dahl and Moretti (2008)

Promissing young women There are several mechanisms explaining why single mothers are so at risk of poverty and why the gender penalty is so strong between single fathers and mothers. To be short and review evidence closer from our empirical setting, we only include research using French data. First, single parenthood is often a consequence of cumulative vulnerabilities. In particular, there is a considerable education gap: 45% of single mothers did not complete high school, 33% hold a university degree, while 50% of mothers in *traditional* households hold a university diploma and 29% less than a high school degree (Le Pape and Helfter 2023, 36–39).

Second, the child penalty is more severe for mothers with low education levels and those at the bottom of the income distribution (Meurs and Pora 2019; Bazen, Xavier, and Périvier 2022). Fast return to work after birth strongly depends on previous labour experience, firm characteristics and the business cycle (Rodrigues and Vergnat 2019). Childrearing also triggers marital specialisation, generating a more inegalitarian gender division of house and care chores, especially among parents with lower education (Bianchi et al. 2014; Solera and Mencarini 2018; Reich-Stiebert, Froehlich, and Voltmer 2023; Briselli and Gonzalez 2023). French women have access to extensive State-provided childcare, rather generous parental leave, and home care allowances that support parental care for two or more young children. Family policies are promoted as giving women a "free choice" and while they are associated with relatively high employment levels for mothers, as a whole, these policies have encouraged women's caregiving rather than promoted men's equal role in care (Misra, Moller, and Budig 2007). Despite being better educated, women generally earn less than men, and continue to shoulder the majority of domestic work even when holding employment (Champagne, Pailhé, and Solaz 2015b; Kandil, Périvier, and Richou 2021). Solera and Mencarini (2018) show that around the first child birth, the traditional division of household tasks, in which women perform more than 75% of all household tasks, increases from 21% to 29%.

Third, by opting for a more family-oriented career, mothers contribute less to the household's income and generally have lower ability to accumulate wealth and savings of their own. Legal statuses of couples (cohabitation, civil union, and marriage) and property regimes (community and separate property) create different levels of protection and compensating mechanisms (alimonies) associated with the gender gap in wealth and saving within couples. As Frémeaux and Leturcq (2022) show, married couples with separate property regimes accumulate more wealth and also have the highest gender wealth gap. Therefore, the single-parenthood penalty is higher among initially richer households while social transfers only partly buffer the shock for low-income households (Bonnet, Montaignac, and Solaz 2024).

Fourth, women's living standards generally decrease more than men's following divorce, due in large part to pre-divorce within-couple earnings inequality resulting from marital specialisation (Bonnet, Garbinti, and Solaz 2021). In addition, mothers are more likely to move than fathers, and shared custody arrangements result in greater mobility for mothers compared to fathers (Ferrari, Bonnet, and Solaz 2019).

Fifth, single-parenthood is not a permanent state and 50% of them are no longer single-parents after 3 years (Costemalle 2017). However, the duration spells vary significantly, with separated or widowed parents leaving single parenthood more quickly than those who had a child outside of a relationship. Most men find new relationships fast while a large share of single mothers remain in this situation.

This "flip side of marital specialisation" leads to a significant re-entry of formerly inactive mothers into the labour market (Bonnet, Garbinti, and Solaz 2021). Shared custodial arrangement have large positive effects on women labour market participation, earning and standard of living (Bonnet, Garbinti, and Solaz 2022). However, Gobillon, Meurs, and Roux (2015) explore gender heterogeneity in access to jobs in France and showed that women have significantly lower access to high-paid jobs than to low-paid jobs. Those who re-partner quickly experience a much lower loss of living standard (Abbas and Garbinti 2019). Bonnet, Montaignac, and Solaz (2024) show that the decline in living standards disappears when the custodial parent re-partners, but this effect only applies to 30% of children six years after separation. On average, children's living standards are 20% lower six years after separation compared to a counterfactual scenario derived from a matched event-study design. Separation also reduces children's educational attainments, more severely for boys than girls and when the separation occurs when the child was young (Le Forner 2020).

III Activating single parents on welfare: Policy environment and intervention

There are 2.25 million single parent households in France in 2018, 4 single mothers for every single fathers. Break-ups represent 78% of entry flows, 16% come from children born among non-cohabiting parents, either by choice or through unplanned pregnancy, and 6% from one parent's death[18]. These families are 2.6 times more likely to live below the poverty thresholds than *traditional* couples[19], with an average poverty rate of 40% ; 22.7% when the custodial parent works and 77.7% among single parents with no job (Le Pape and Helfter 2023, 36–39). Lone parents and single-parent families are particularly vulnerable to situations of insecurity and poverty, and they tend to have a pessimistic view of their current situation, their future, and society as a whole (Pirus 2021). They are also strongly in favour of more generous support for families. A large share live in poverty, relying heavily on various social transfers.

In this section, we first start by presenting what social transfers they may receive and take the example of a single parent family with two children to illustrate their interactions. Then, subsection 2 presents the **Reliance** programme, a randomised intensive welfare-to-work programme in place from 2018 to 2022 in the North-East of France. We discuss the results of the first evaluations focusing on the first causal evaluation of Heim (2024) and the qualitative study of FORS (2020). From there, we motivates our analysis with a discussion on the objective of the programme, some key components of the intervention and suggestive evidence from the qualitative evaluation. Finally, we discuss the incentives of the tax-benefit system in subsection 3 and the possible behaviours that participants may exhibit when these incentives are made salient.

1 Social transfers for low income families

The French tax-benefit system involves many different policy instruments targeting different populations. In Appendix A, we review the history and the evaluation of policies introducing monetary incentives in France. In short, low income families have access to different aids which serve different policy objectives. The main ones are:

1) **Minimum income scheme** RSA (*Revenu de soliarité active*) and **in-work benefits** PA (*Prime d'activité*). They are mean-tested monthly payments based on household's quarterly income and family structure;
2) **Housing benefits** AL (*Allocations logement*) depend on rent and number of dependent children, but not on cohabitation. They depend on household's income from the past 12 months as a moving average updated quarterly[20];
3) **Family benefits** AF (*Allocation familiales*) for parents of two, with mean-tested **supplements** CF (*Complément familale*) for families of three or more children. There is also a mean-tested **back-to-school** allowance paid in summer;
4) **Early childhood benefits** PAJE (*Prestation d'accueil du jeune enfant*) which includes a mean-tested basic allowance: AB (*Allocation de base*) and **childcare allowance** CMG (*complément mode de garde*) for active parents resorting to childminders, and **shared parental leave** PrePare (*Prestation partagée d'éducation de l'enfant*);
5) **Disability benefits** AAH (*Allocation adulte handicapé*) for parents, AEEH (*Allocation d'éducation d'enfant handicapé*) for children, and AJPP (*Allocation journalière de présence parentale*) when parents take care of a sick child.

Parental leave provides between 50 and 2/3 full-time minimum wage but is barely used by single parents because this level does not compensate for the time off work, whereas mothers in couples have the resources of a spouse (Périvier 2022a). Those transfers are not specific to single parents but take their situation into account either through

[18] Costemalle (2017) based on INSEE data from the Family and Housing survey 2011.

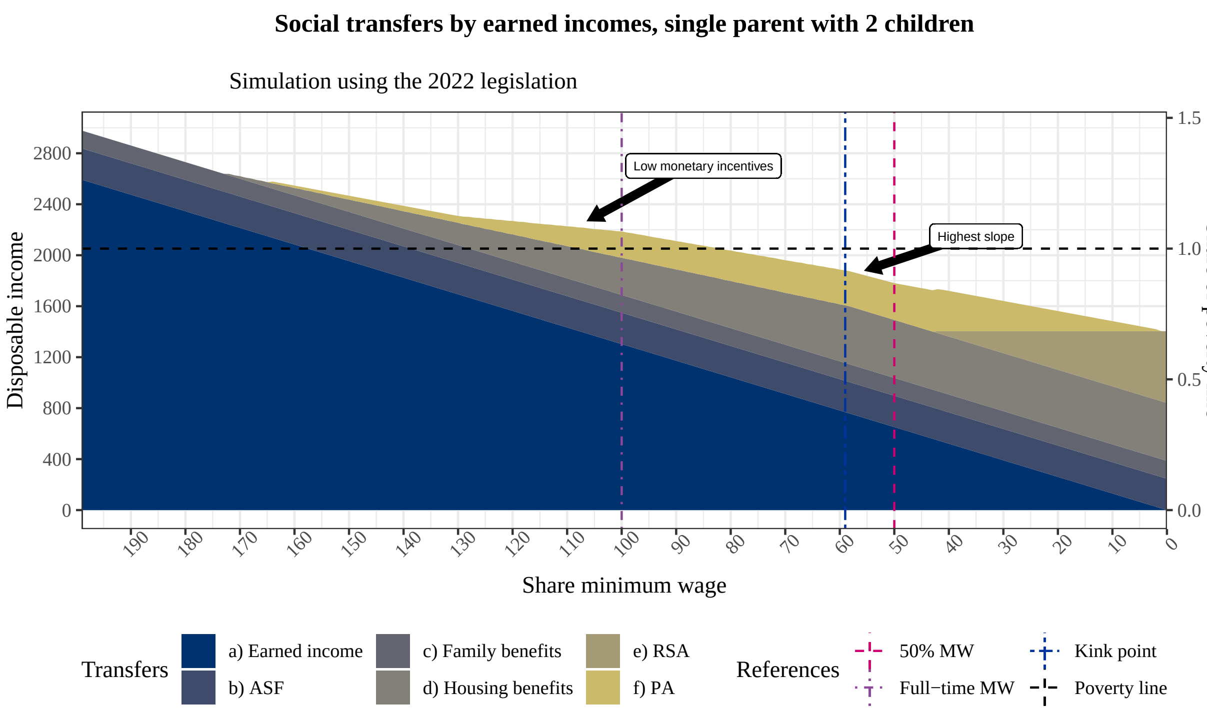
[19] The poverty rate among *traditional* couple is 15.6%, 4.7% in dual-earner couples, 30% among single-earners and 71% when no parent works.

[20] There are actually three types of housing benefits: APL (*aide personalisée au logement*) for anyone that rents a house, which is only contingent upon income and rent prices. ALF (*Allocation de logement familial*) is open to households with family allowance or disabled child allowance but also married couples for the first 5 years after marriage. ALS (*Allocation de logement social*) for other cases.

increased amounts or eligibility cut-offs. The only specific aid is the **family support allowance** ASF (*Allocation de soutien familial*), a lump-sum quarterly transfer for custodial single parents who receive no or too little child support from the child(ren)'s other parent. Its amount depends on the number of children and is first offered for four months. To keep receiving ASF, single parents must prove they started the administrative or judiciary process to set child support. If the other parent cannot pay child support, the administration verifies and can condition ASF to new judiciary processes. If child support has been defined but is not (entirely) paid, ASF is an advance and the administration will try to recover the sum from the other parent. Any child support paid is deduced from ASF, and ASF is mostly deduced from RSA and PA. Last, if the single parent starts living with a partner again, they are no longer eligible for ASF. We come back to these points in the subsequent paragraphs. In November 2022, its amount has been increased by 50%, going from € 122 to € 184 top per child.

Except for pregnancies which are reported directly by physicians to the health system, any other change must be reported by households as it defines many eligibility rules and amounts of cash transfers. Failure to report changes can have severe consequences for households. In case of *undue payments*, they may be required to repay the amount for up to 2 years back.

Figure III.3: The static model of the French tax-benefit system for a single parent of two.



Sources: DREES, EDIFIS.
Case study for single parent with two children reiving RSA, no temporary supplement.
Arrows and labels indicate pieces where the implicit marginal tax rate is locally highest or lowest.
a) Earned income as a share of the full−time minimum wage (MW, 1302.67 euros in 2022).
b) ASF is the public child support paid quarterly when the other parent doesn't. Lump−sum payment by number of children.
c) Family benefits are open for parents with 2 or more children, depend on taxable incomes (Year − 2).
d) Housing benefits, baseline amount depends on number of dependent children, rent price and their due payment.
Decreases with earned incomes over the past 11 months, from the previous month.
e) RSA is the French minimum income scheme, f) PA is an in−work benefit, both are mean−tested
and depend on family composition and household's income.
Poverty line is 1140 euros per Consumption unit (CU) times 1.8 CU (1 for the parent,
.3 per child between 3 and 14 and .2 as a single parent).

To fix ideas, consider a single parent family with two children between 10 and 14 years old and no income in 2022. This parent receives neither temporary supplement-RSA, nor children allowance, but receives ASF for the two children. We use the open-source simulation model of the Statistics department of the Ministry of solidarity[21] to represent all these social transfers and their interaction with labour incomes in Figure III.3.

[21] Accessible from https://drees.shinyapps.io/Drees_Maquette_Edifis/

RSA and PA are differential incomes. They are computed as the difference between a statutory level based on household size and composition from which the relevant *reference income* is deduced, which defines the final amount received. We provide details on the formula and an important reform that took place during our time-frame in Appendix A. There are three important features. First, households must report all their incomes from the previous quarter by household member and type of income. Second, they depend on three measures of income: joint labour earnings over the quarter, individual labour earnings and joint pre-tax incomes. Third, any reported income is deduced (almost) entirely from RSA while PA considers all labour incomes together, with an individual bonus starting at 50% minimum wage[22].

Without labour incomes, a 2-children single parent household receives up to €1492 if they access all their rights. Accounting for family size, this amount to €933 per consumption unit[23] which is 18% lower than the poverty line[24]. Note that the equivalent scale of consumption units tend to overestimate the standard of living of single parents and parents without full custody (H. Martin and Périvier 2018). For single parents in this situation, 38% of their incomes are mean-tested and change every quarter (RSA and PA); 30% are housing benefits and they are now changing monthly using a moving average from income 12 month earlier to two month before the month of payment[25]. The remaining 26% are not mean tested.

Last, this static model shows that the households' budget increases rather smoothly with labour income while the amount of social transfers fades out until 1.8 times the full-time minimum wage (MW hereafter). On average, the implicit marginal tax rate is about 38% but there are three important non-linear pieces. The first notch and following convex kink are due to i) a €15 minimal payment threshold at the RSA exit-point, ii) the starting point of individual[26] PA supplement from 50% MW, and iii) maximum amount of housing benefits up to 60% MW. The second convex kink starts right at the MW level as the amount of PA rapidly decreases up to 1.3 MW, where last piecewise non-linearities create non-convex kinks. The alignment with the full-time minimum wage is a consequence of the 2019 reform of in-work benefits - a point we discuss and illustrate In Appendix A.III.

Single parent households on welfare receive no money from child support. Family benefits are not a significant portion of the budget of a single parent of two on welfare. Nevertheless, ASF is a crucial element of the overall budget and has strict eligibility criteria, as previously mentioned. It is important to note that a single parent without ASF would not lose a proportionate amount. Actually, most[27] of the ASF and 100% of family benefits have already been deducted from the RSA. Conversely, when single parents receive child support from the other parent, this amount is 100% deduced from RSA and PA. Their income does not change when they receive it or not. For these families, child support neither benefit the child nor the parent; rather, it reduces public spending[28].

This feature has important consequences for poor families. First, as noted by Pucci and Périvier (2022), it creates strong inequalities between high and low income single parents, where the former keep ASF or child support in full while the poorest lose all child support or keep only a portion of ASF. Périvier (2022a) shows that because of their inclusion in baseline ressource of RSA and PA, better recovery of unpaid child support can reduce the standard of living of lone parents, particularly lone mothers. This reduces incentives to make the other parent pay since the custodial parents receive nothing or even be poorer. Moreover, child support and ASF also reduce share of mean-tested benefits in all transfers. Consequently, their exit points in the earning distribution also gets lower with any non-labour income.

[22] There is a subtlety in the precise computation: PA is based on each month's labour income over the quarter while RSA takes all incomes in the quarter and average them. One month with more than € 1500 gives PA for one month but remove RSA for one quarter or reduced it for two, depending on which month was worked in the quarter.

[23] The CAF uses the following weighting: 1 consumption unit for the first adult, 0.5 for each adult or child aged 14 or older, 0.3 for children under 14, 0.2 for single-parent families. For a discussion on the limits of equivalent scales, see H. Martin and Périvier (2018).

[24] € 1140 in 2019.

[25] This new computation formula has been in place since January 1st, 2021. Before, the reference income was the monthly average of the taxable income, therefore based on two-years old incomes.

[26] The PA formula incorporates an *individual bonus* that commences at a minimum of 50% of the wage, which serves to diminish the marginal tax rate while simultaneously taking into account the number of individuals in households with earnings exceeding the designated threshold. However, it should be noted that these individual bonuses are not awarded separately to each contributing member; rather, they contribute to the overall amount. See Appendix A.III for details.

[27] Before the 2021 reform, 80% of ASF was deduced from RSA and PA. The increased has not been changed in RSA and the share deduced is now 53% (DREES 2022).

[28] For a discussion and proposition of new policies, see Pucci and Périvier (2022).

social transfers are tightly linked together but use different notions of income. The smoothness of the budget constraint largely depends on the stack of social transfers parents are entitled to, but these transfers have different eligibility criteria. They use different measures of income, adjusting more or less rapidly to changes and accounting for family structure differently. For instance, RSA and PA are adjusted based on household's quarterly earned income and family structure, while housing benefits use a moving average of earned income in the past 12 months and depend on the number of children, rent, and household income. They are the same amount with or without a partner but use the same formula with one or two incomes. The other transfers are either not mean-tested (ASF) or based on tax-income from 2 years before. As a result, social transfers do not adjust to changes in labour income or family situation in the same way. In Appendix A.II, we demonstrate how challenging it can be to track and understand changes in social transfers in a simple scenario of a single-parent taking a job.

In December 2018, 76% of RSA recipients had been receiving it for more than a year, with an average registration time of 5.9 years, and 59% received it continuously between 2008 and 2018 (DREES 2022). **Total cash transfers are lower than the poverty line and RSA recipients often remains in this situation for many years**. The French minimum income (RSA) is conditioned upon participating in mandatory social support and job search. However, their implementation have been strongly criticised in official reports (Pitollat and Klein 2018, Aout ; Damon 2018 ; Cour des comptes 2022). In a 2022 survey, Athari (2023) reports that 45 % or RSA recipients received no support in the past 12 months, 87 % are involved in a social programme but they start late: only 59 % of those registered for less than 6 months are.

2 Reliance: A randomised intensive welfare-to-work programme

From 2018 to 2022, we ran a randomised control trial of an intensive support programme in the North East of France. This high stake programme has been supported by the National family allowance fund, the local administration in charge of workfare obligations, the Employment agency and a public investment fund (*caisse des dépôts et consignation*) and cited as a benchmark for future welfare-to-work programmes in the aforementioned reports. The assumptions regarding the effects of the programme are based on both a *capacity-building* approach – where participants benefit by developing or maintaining skills, building a network, etc. – and an *emancipatory* approach that seeks to alleviate the specific burdens and obstacles faced by each family. Both reflect the '*social investment*' perspective of this programme, notably inspired by the "*capability*" approach of social justice from Amartya Sen (Martignani 2016; A. Hemerijck 2018). The average cost per participant is estimated to be around €2800, approximately four times the average expenditure for regular support (Mahdi 2021). However, the recruiting process was more embedded in the *workfare* approach of active labour market policies.

A staggered block-randomised encouragement design This programme is based on a staggered block-randomised encouragement design. Each of the five years of implementation, a random sample of 500 eligible households has been drawn from administrative records. Social workers from the Departmental council assessed and classified every file and those already in another programme, in employment, or deemed inapt, were excluded from the experimental sample (1/5 of the initial sample on average). The remaining households in each cohort were then randomly assigned encouragement within blocks based on the cross product of number of children (1, 2, 3+), registration at the Employment agency (True/False) and number of years on welfare (2-5, 5-10, more than 10 years). The product set of these variables defines single parents' *type* between which we expect different reaction to the encouragement, different outcomes and possibly heterogeneous treatment effects. Registered unemployed are expected to be closer to the labour market, the number of children increases constraints due to parental obligations while it has been shown that the longer people receive RSA, the less likely they are to find a job[29].

To ensure high levels of compliance, we adapted our recruitment process over the years, employing both "threats" and "pull" strategies in our recruitment process. Initially, we sent a formal letter to parents inviting them to a public meeting in the welfare-to-work department of the Departmental council. The following year, we added an ambiguous yet threatening sentence about "rights and duties" in the invitation letter, which served as a "threat effect" to encourage participation. Additionally, we made our recruitment sessions more welcoming and personalised by moving from collective information sessions in the welfare-to-work division of the Departmental council to

[29] See Heim (2024) for more details and descriptive statistics across groups

individual face-to-face interviews with project managers in newly renovated and well-equipped premises. We also involved former participants in the recruitment sessions to answer questions and provide positive testimonies as peers. Table III.1 in the Appendix summarises the average effects of encouragement on participation and separates estimations by cohort[30]. On average, we achieved a take-up of 38.9 in the four first cohorts, increasing by 19.25 pp between the first and fourth.

Year-long social support, childcare and job-search assistance The initial stages of the programme focused on tackling major issues such as over-indebtedness, housing, healthcare, and children's education. The programme was designed to be conveniently scheduled in relation to school and daycare timetables, and participants were required to commit approximately 15-20 hours per week. Childcare duties were shared among participants in a dedicated, colourful space that was equipped with baby-care supplies, toys, games, and books. The programme's target population required a lot of social support, and the *Reliance* programme combined "classic" individual support with group support through thematic workshops. Activities primarily focused on creating and validating realistic professional projects, addressing steps like education, internships, and improving job search efficiency, aligning expectations with job opportunities. Some workshops addressed daily life and organisation, offering strategies to cope with upcoming changes and find appropriate solutions. Others explored self-awareness and relationships with others, including issues related to parenthood, relationships, and gender norms and roles. The programme aimed to humanise administrative procedures and alleviate the emotional burden associated with them. For instance, social workers offered assistance with applications for social housing, affordable school lunch or suitable childcare options. Lastly, the programme organised regular sessions with social workers from the Family allowance fund. These workshops focused on access to benefit rights and parenthood to help participants understand and know how to gain access to their rights. We will return to this important point later.

Policy and macro-economic changes in the timeframe During this period, the economic environment was significantly impacted by a major and sudden reform of PA in 2019 and the COVID-19 crisis from March 2020.

First, the pandemic disrupted the implementation of the programme for the 2020 cohort, and the economy was almost entirely shut down for a while before slowly recovering.[31] As a result, we decided to continue with the programme and secured funding for two additional cohorts. To increase precision, we enlarged the sample for the 2021 cohort. In 2022, the pool of eligible families that had not been sampled was too small, especially for long-term recipients. To build the 2022 cohort, we sampled the remaining eligible population and random samples from the control groups of the previous cohorts. From November 2021, the composition of the control groups of the four previous cohorts changed. Ultimately, the intervention was rolled out from 2018 to 2022 in a staggered design, as illustrated in Figure A.11 in the Appendix.

Second, the 2019 reform of in-work benefits was adopted unexpectedly as an answer to massive demonstrations from the Yellow Vest movement. In the timeline of this experiment, it occurred at the last quarter of the first cohort and right when the second was being recruited. At that time, we had meetings with project managers to discuss how to explain the reform to participants and further emphasise monetary incentives in the programme. We provided simple plots[32] of the amounts of social benefits over incomes in percentage of the minimum wage from simulation models of the Family allowance fund (A very similar plot as Figure III.3). We also provided figures illustrating the effects of the reform, borrowing results from colleagues who simulated the reform to estimate the cost of the reform ex-ante[33]. Unbeknownst to us for years, social workers used these tools a lot and we only discovered that when they asked us for an updated version during a meeting.

[30] Simply regressing participation on encouragement and block fixed effects.

[31] For more details on the adaptation and consequences of the pandemic, see Appendix A.

[32] Figure III.3 is one example although this one uses the simulation model from DREES while we used the one from Cnaf at that time. These models are almost identical by design.

[33] Results based on these estimations have now been published by Dardier, Doan, and Lhermet (2022), and the figure we mentioned is reported in Figure A.18 in the Appendix.

No effect on employment and disposable incomes after training In a companion paper, Heim (2024) offers a first examination of the programme's effects and provides further details on implementation and participant characteristics. The findings indicate that the programme slows the job finding rate during its initial phase, leading to a pronounced lock-in effect on poverty rates, disposable income, and employment. However, these anticipated negative effects wane by the end of the programme, and there are no average effect for the year after. The central message of this paper is that without random assignment, one would incorrectly conclude that the programme enhances employment. In reality, the programme attracts individuals with the highest potential employment levels but does not boost labour market participation. The selection bias is so strong that estimates using modern doubly-robust matched difference-in-differences fail to include the experimental results within the simultaneous 95% confidence interval. This research also showed that participants were more likely to be among the poorest and least educated but also more likely to be registered at the Employment agency, a group notably more likely to find a job both among participants and in the control group.

An independent team of consultants also conducted a qualitative evaluation from 2019 to 2020 (FORS 2020), adding the views of participants to the analysis of the implementation. They note that the initial three months of the programme creates a rapid shift in momentum, boosting confidence and self-esteem relatively quickly. The group dynamic and "lever effect" of collective activities helped individuals overcome isolation. However, isolated individuals with limited work experience and reduced autonomy, and those with active social networks but not actively seeking employment due to care-giving responsibilities, health issues, etc. faced a longer timeline for employment reintegration, as their focus was not on immediate employment but rather on formulating a professional plan.

The initial findings indicate that this programme had a strong *screening effect*, drawing in individuals who were most likely to secure employment. These results align with the qualitative evaluation's findings. However, the programme did not prove effective in expanding employment opportunities at the extensive margin. After 30 months, 89% of participants remained in poverty and 66 % have no job.
The analysis of the heterogeneity of treatment effects in Heim (2024) revealed puzzling patterns, particularly by median income and number of children at baseline. The effects on disposable income and labour market participation exhibit divergent patterns, implying changes in the composition of income, which is corroborated by the rise in cash transfers at the end of the period. This can only mean that the programme affected family composition and source of income differently for parents with more or less children at baseline.

The problem with this analysis is that we have very little guidance for interpreting the results. However, welfare and policy implications are not the same if the programme affects household size through fertility, the merging of two single-parent families, or if older children remain or leave the household. In Heim (2024), the research was guided by the policymakers' objectives and a typical evaluation problem. This helps to define clear normative criteria and decision rules. However, it cannot explain *why* the programme did not increase employment beyond selection effects. To provide more informative results on the mechanisms, we briefly reopen the black box of the programme with insights from the qualitative evaluation.

As A. Hemerijck and Huguenot-Noël (2022) write, "*The normative heart of the social investment paradigm, beyond fair distribution, underlines the importance of secure capabilities enabling citizens to flourish over the life course*". In line with this view, favourable outcomes may not be a matter of resources or preference satisfaction but, rather, of what a person is able to do and be. As such, poverty could be understood as agency-based capability deprivation (A. Mani et al. 2013; Shah et al. 2018; Banerjee, Duflo, and Sharma decembre 2021; Henderson and Follett 2022).

In this setting, participants spent a large amount of time in this programme and what we know from the qualitative evaluation is that many participants really felt that the programme helped; at least those interviewed. The qualitative evaluation is filled with verbatims of "empowered single mothers" with higher sense of agency and expanding horizons.

Optimisation as functionnings The programme's interaction with participants through regular visits from social workers provided valuable opportunities for knowledge-building and empowerment. These interventions ensured that participants understood their rights and the complexities of eligibility criteria and benefit calculations. By using visual aids such as Figure III.3 of the tax-benefit we provided, social workers illustrated to participants that pursuing employment would not result in a complete loss of allowances. This hands-on approach was instrumental in equipping participants with a deeper understanding of their rights and future prospects.

As highlighted in the qualitative evaluation by FORS (2020), simulations of returning to work organised with the CAF were particularly enlightening for participants, enabling them to envision their future financial resources. Testimonials express newfound awareness of their entitlements and a shift in mindset towards employment. For instance, one participant reflected, "*I know I have rights, but at the moment I can't necessarily open them (alimony, housing benefits, etc.), I asked at the CAF when they came. That's what I learnt at Reliance, that I could activate that. We found out about CAF-related stuff, etc., and that's something*" (F, 41 years old, 3 children, separated for 2 years).

Another participant shared, "*When I first came here, I remember a discussion with [Project manager] where we talked about income: I didn't want to work more than half-time either, so as not to lose income, and she told me I was wrong. And in fact she wasn't wrong: my attitude did in fact change. That day there was a social worker from the CAF with us, and I was stubborn and narrow-minded*" (F, 32, 2 children).

These anecdotal evidences suggest that some participants benefited from the programme by learning about French tax-benefits, rights they were previously unaware of, or eligibility criteria. However, it also highlights that some participants' understanding of their rights remained limited despite years in the system. Nevertheless, the programme's use of visual aids, simulations, and personalised advice may have contributed to the development of enhanced agency and capability.

If these results go beyond anecdotal evidence, we expect participants to have more *rational* reactions to salient incentives. If these incentives are strong, they can generate reactions large enough for us to measure.

3 '*Assistaxation*': economic incentives and family structure

We want to know what the monetary incentives are for single parents on long-term welfare and understand how it differs for single/coupled parents with different numbers of children. Our approach consists simply in visually inspecting plots of social transfers when we vary the number of children and cohabitation with labour incomes of a single earner. We use the EDIFIS model to simulate these situations and report the results in Figure III.4.

Three facts on welfare benefits for single parents in France: This graphics shows three under-reported stylised facts of the French benefit system:

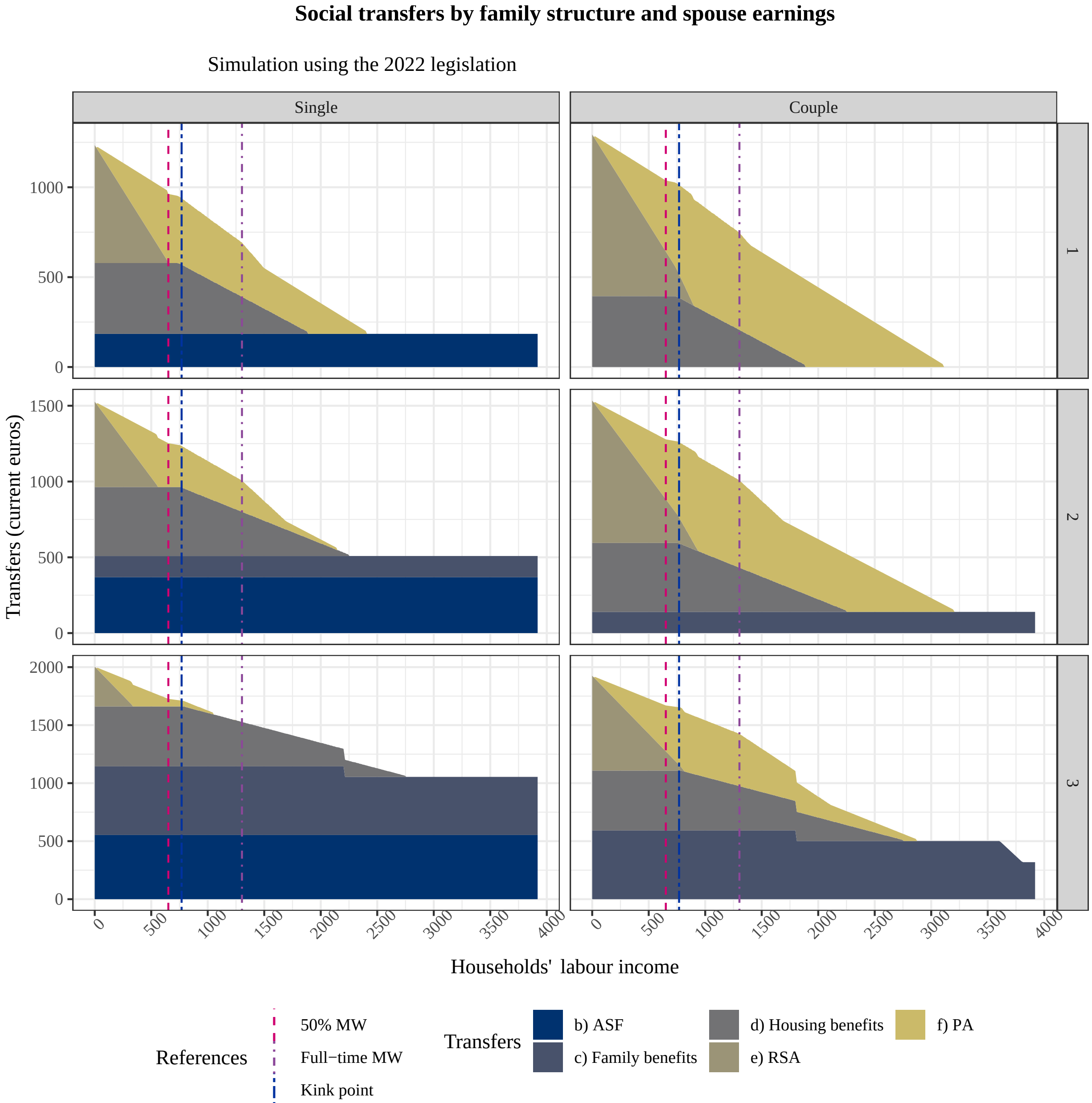
Fact 2. *The composition of total social transfers is very different across household' structures.*

Fact 3. *In-work benefits decrease at faster rates after the minimum wage for single mothers than for couples. Its amount are much smaller and exhaust more rapidly than for couples.*

Fact 4. *There is a range between .5 and .6 full-time minimum wage with minimal implicit tax rate and stable levels of all social transfers.*

All three have important consequences and incentives. With Fact 2, we see that almost all incomes for single parents with one child are mean-tested, but housing benefits do not depend on cohabitation and adjust more slowly. For single parents with two children, family benefits and ASF represent an important part of the budget and are largely deduced from RSA and PA. Therefore, in-work benefits and RSA are lower and exhaust faster. Note that if the parent was receiving child support instead of ASF, there would be an even larger share deduced from RSA and PA. Because of this feature, parents of three receive so little RSA and PA that they do not receive in-work benefits with earnings over 3/4 of the full-time minimum wage. Receiving child support makes it even worse. For single parents of three, their level of transfers depends mostly on their number of dependent children for both housing and family benefits.

Figure III.4: Monetary incentives for households with 1 to 3 children, with or without a partner (in which case, x axis is household income)



Sources: DREES, EDIFIS.
Case study for single parent or couple with only one varying labour income.
b) ASF is the public child support paid quarterly when the other parent doesn't. Lump−sum payment by number of children.
c) Family benefits are open for parents with 2 or more children, depend on taxable incomes (Year − 2).
d) Housing benefits, baseline amount depends on number of dependent children, rent price and their due payment.
Decreases with earned incomes over the past 11 months, from the previous month.
e) RSA is the French minimum income scheme, f) PA is an in−work benefit, both are mean−tested
and depend on family composition and household's income.

Fact 3 has strong political and fairness implications. The overall interactions of all allowances and complex eligibility rules make them very hard to understand and can create large variations on the implicit marginal tax rates, as Figure A.19 in the Appendix illustrates. Using again the simulations from the EDIFIS model, we report implicit marginal tax rates in variation with earned incomes by percent of minimum incomes. These interactions create different plateaus with highly different tax rates. Consistent with these figures, the highest marginal tax rate is located at the minimum wage level for single parents with one and two children. Between 100 and 110% of the minimum wage, the IMTR is 72% for single parents of one child and 69% between 100% and 120% of the minimum wage for single parents of two children. These variations of implicit marginal tax rate motivate our analysis of the reaction at the intensive margin using bunching around kink points to retrieve observable elasticities. In Section 1, table VI.1 reports the average implicit marginal tax rates, giving a clearer picture of the sharp inequalities and incentives.

The most typical labour contract for low-educated workers is the most heavily taxed for single parents. It is also the lowest level of payroll taxes (Bozio, Breda, and Guillot 2023). This makes it much harder for single parents to increase their disposable income by working longer hours. This is particularly the case for single parents with two children, for whom any improvement from the full-time minimum wage results in a rapid and sharp reduction of in work benefit: a €100 monthly increase at the minimum wage level induces a reduction of in-work benefits of €75 per month in the next quarter. Having higher constraints than those with one child (Briard 2020), **this group has the largest disincentive to work-full-time**.

The situations is very different with a partner, where a single wage opens large amounts of in-work benefits, even with one part-time job. For single parents with one child, finding a partner with a child would compensate the loss of ASF through family benefits. For parents of two, however, cohabitation means losing €246 of ASF.

We are not the first to identify the unfairness and sharp inequalities between single parents and more traditional families. Périvier (2012) offers additional arguments, and notably that mandatory support and job search for RSA recipients is lifted when households earn more than € 500. Couples in the traditional bread-earner models are not *bothered* by workfare obligations while a single mothers receiving RSA is, which the author sums-up in the article's title: "*work, or get married*". Paradoxically, while single parents are most heavily taxed on full-time jobs, statistical data from the Labour Force survey indicates that 40% of those working part-time express a willingness to work more hours if given the opportunity. Interestingly, partnered women in part-time positions are only 22% as likely to express this desire (Périvier 2022a).

Hidden tax on single parents: the assistaxation trap. In the qualitative study of FORS (2020), several individuals and group interviews revealed that, in most cases, ex-partners are disengaged or even absent in the upbringing of their child(ren), and their financial contribution to material expenses is often minimal or non-existent. At Baseline, only 20.6% of the sample receive child support and therefore, 65 perceive ASF instead[34]. However, some mothers express reluctance to apply for the ASF, fearing the repercussions from their ex-partner or potentially harming the quality of the relationship that the ex-partner maintains with their shared child(ren). "*A while ago, I applied for ASF; the dad and I didn't get along. We went to court, and he was supposed to pay child support, but he never did. But now that we get along a bit better, and when the girls need something, he still buys it for them. We prefer not to reapply. Otherwise, I'm afraid we won't get along anymore.*" (A., 40 years old, 2 children, on RSA for over 10 years). Pucci and Périvier (2022) recommend to remove ASF from the reference incomes for RSA, PA and housing benefit and to neutralise at least part of child support so that single parents who receive it actually see their income increase.

As for parents of three children (or more), the inclusion of family benefits in the baseline income for RSA and PA drastically reduce their amount so much that the exit point of PA is below 80% minimum wage. That is because having 3 children or more increases family benefits and (may) open an additional allowance[35] that are both immediately and entirely deduced from RSA and PA base level. For single parents of 3 or more children, RSA and PA do not play the role of buffer and monetary incentives, these are done via family and housing benefits. Paradoxically, it also means that their main source of income is 100% taxed until they leave welfare and have a lower and flatter implicit tax rate. We also note that family and housing benefits are not taxable incomes. They are only fully taxed for those receiving social assistance.

According to Fact 4, individuals in all three groups find a *sweet spot* in the labour earnings distribution between 50% and 60% of the full-time minimum wage level. This holds true even for those who repartner. For parents of two children, recognising the taxing nature of the full-time minimum wage system should be highly discouraging.

[34] The remaining share receives neither.
[35] *Complément familal*

Part-time jobs also serve as a strong reference point (Tversky and Kahneman 1974). They represent an attainable goal that may have been encouraged by social workers. For households with a strong aversion to loss, part-time jobs do not impact housing benefits, only RSA and PA adjust. However, such jobs usually do not generate income high enough to get out of poverty and often have limited opportunity to build more human capital (Blundell et al. 2016).

The complex interactions between various social transfers with differing schedules result in an increased tax burden on one of society's most vulnerable groups. We propose that "*Assistaxation*" is an appropriate term to convey the concept of providing assistance that becomes burdensome, overly taxing, or difficult mentally, physically, emotionally, or financially. It is not just about the administrative burden and stigma discussed in the literature; it also involves heavier and implicit taxation of labour income and a 100% tax rate on child support and family benefits of the poorest.

At this point, if the programme allowed participants to make better inferences about their expected income, few reactions to *assistaxation* could follow:

1. Part-time instead of full-time work, with a focus on the *sweet spot*.
2. Single parents with one child have little to lose by repartnering and may even receive higher social support if their partner also has children.
3. Those with two children are strongly discouraged from working full-time at the minimum wage and above, with no clear incentives regarding cohabitation.
4. The income of those with three or more children mostly depends on family benefits, and monetary incentives are limited.

We now present the data we use to investigate such behavioural changes following this experimental welfare-to-work programme.

IV Data and descriptive statistics

The design of this experiment relies on multiple random samples of eligible households extracted from the administrative records of the Departmental council. These samples are organised into cohorts that have been randomly assigned treatment following the protocol outlined in the previous section. The initial datasets have been enhanced with these design variables and aligned with the monthly administrative records from the National family allowance fund (CNAF).

1 Source and definition of the main variables

The ALLSTAT files provide information on the situation of every beneficiary for a specific month. They include details on the "household heads" and their possible spouse, such as gender, year of birth, marital status, activity status, nationality, and more. Additionally, they provide information on dependent children, including their year of birth, alternate residence status for family benefits, absence of a parent, the legal benefits they receive, and various measures of household income. The complete database spans from January, 2017 to June, 2023. Cohorts have been randomised at different time and we define the month before random assignment as the main reference to measure the average dynamics of treatment effects across cohorts. To avoid bias, we restrict our window of observation over the 30 months after random assignment over which all four cohorts are observed. We provide details on the pre-treatment of our database and balance check in Appendix B.

In this research, our main variable of interest are incomes and family composition for which we have several variables.

Measures of income We use self-declared incomes of the quarterly report for RSA and PA to define a set of measures for individual and household incomes[36]. In particular, we define four measures of pre-tax/benefit incomes:

- **Individual's labour incomes** are the labour incomes earned by the participant of the programme reported to the administration ;
- **Individual's pre-tax incomes** are labour incomes and other sources of income such as child support or other non-labour incomes.
- **Household's labour incomes** are the sum of all labour incomes reported. It includes potential spouse labour earnings.
- **Household's pre-tax incomes** are the total household income before tax and benefits are adjusted.

The database also contains **total cash transfers** and with it, a measure of **household disposable income** which is the post-tax-benefit income of the household. Then, the Family allowance fund defines the **number of consumption units** and computes the **disposable income per capita** accounting for family size. The number of consumption units is 1 for the first adult, 0.5 for each adult or child aged 14 or older, 0.3 for children under 14, 0.2 for single-parent families. This latter variable can serve as a composite index of family structure over which parents have agency.

All monetary values are converted in constant € in 2015 values using the consumer price index of the bottom 20% of the income distributions. We also use the full-time minimum wage as a reference point and use variables divided by the value of the actualised minimum wage that month.

Couple formation To measure if the programme affects family composition, we consider **cohabitation** with a partner as a first outcome. For that, we use the binary variable from the ALLSTAT files that codes 1 if the person lives with another adult with whom she is in a romantic relationship[37]. This variable is very important and conditions most payments. It is either modified after the parent reports her change or a control from CAF. It captures living together officially rather than romantic relationships *per se*.

Cohabitation is a very confuse notion. For the Family allowance fund, cohabitation is made of two people living together, considered as a couple by their social circle, who share financial responsibilities and household duties. However, living separately from one's partner can still be considered cohabitation if the two share financial responsibilities. Conversely, a room-mate is not a cohabitation unless they are known to be romantically involved.

This variable captures official cohabitation which may be affected by the programme through increased share of relationships - possibly with a former partner - and higher reporting among participants with better knowledge of the definition and stakes. Cohabitation can be more or less incentivising for parents, depending on their situation and that of their partner. For the Family allowance fund, it means entirely pooling their income and possibly including other children.

Children number and characteristics There are several measurements for the number of children in the ALLSTAT database, because they are used for different aids and some have different eligibility rules. Households receive family benefits until the child is 20, but housing benefits and additional family benefits can be extended until the child is 21. To be consistent with other outcomes, we use the variable updated quarterly with income report for RSA and PA and use the others as robustness checks. We construct variables of the oldest child in a given month and use it to define quartiles of households with children close to autonomy at the time of random assignment. The top quartile of that variable corresponds to 17 years old. This means that parents in this quartile have children old enough to move-out by the time the programme ends and soon after.

We measure fertility with the probability that the mother is pregnant a given month. The variable is a dummy coming from social security flows after a doctor declared the pregnancy. This very reliable measurement allows to capture fertility effects about 6 months before birth, which may be important considering the censored window of observations.

[36] For administrative clarity, these monthly gross incomes are separately identified for "*Monsieur*" and "*Madame*" (as per the terminology still used in the CNAF administrative files 12 years after the same-sex marriage law was passed). The constructed variables are defined accordingly.
[37] And compute the eligibility and amounts of all transfers accordingly.

Timing of events We want to investigate how taking a job affects our final outcomes and possibly other dimensions. For that, we want to compare changes around this event. We define a pair of variables for the **month-of-event**, and **time-to-event**, which measures the number of months relative to the month of event. Month-of-event defines groups of single parents for whom the first job re-entry occurred the same month. We explain how we use these variables in section V.

2 Descriptive evidence of important heterogenous reactions

Following our hypothesis, we first look at the distribution of reported earned incomes across treatment groups. We first start with household's labour income which is the main income used to compute PA.

Bunching at the *sweet spot* Using random assignment as a reference point, cohorts started the programme from the 2^{nd} to the 6^{th} month after and the programme ended between 14 and 18 months after random assignment. We look at the 12 months from the 18^{th} to the 30^{th} since random assignment and only look at observations with positive labour income. Figure IV.5 presents the kernel-density estimates of households' labour income, separated by actual participation and number of children at baseline. The in-work benefits are computed from this variable. In line with the optimisation friction and bunching literature, we also added the theoretical amount of in-work benefits (PA) using the EDIFIS simulation model, and the *sweet spot* where the implicit marginal tax rate is lowest and housing benefits do not decrease.

The distribution of labour incomes among treated households contrasts sharply with the other groups, who overlap for the most part. Participants with one or two children report a mass density of household's labour incomes within the range of €600_{2015} to €800_{2015} net per month; the *sweet spot* previously identified. Conversely, control and never-takers' distributions have thicker tails than participants. There is a thicker distribution of part-time jobs among never-takers with two children than in the control group. The income distribution for single parents of three or more is higher for part-time jobs, but similar across groups, with thicker and longer tails. Figure C.21 in the Appendix displays the same kernel densities over individual labour incomes and shows a very similar pattern.

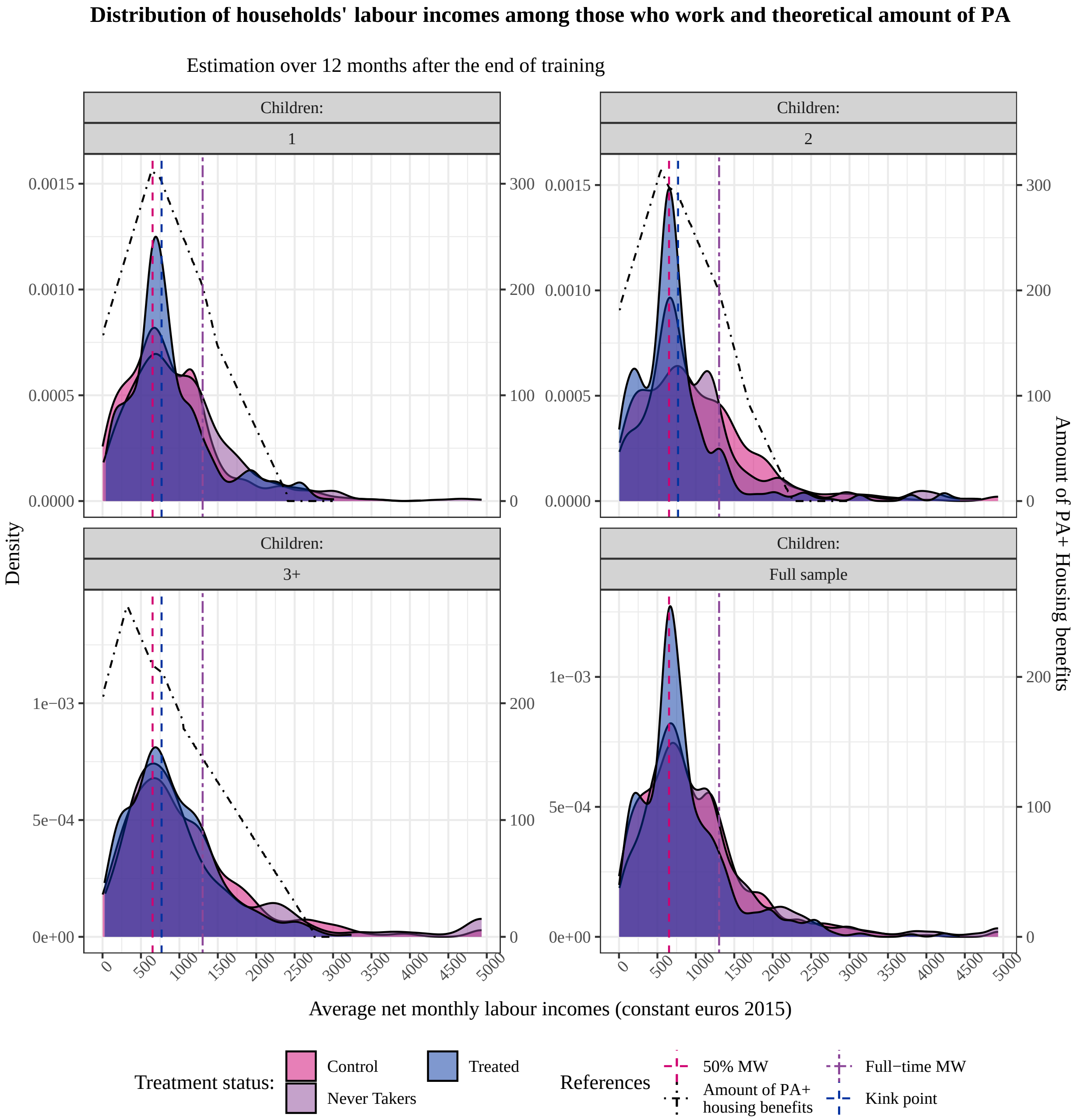
Without any assumption, these plots show that participants (or their partners) are more likely to work part-time than other comparison groups. However, this observation is not enough to make a compelling case on participants capabilities. It may very well stem from advises from social workers and the definition of part-time job as an objective. However, PA is computed over the joint-distribution of labour income for all members of the household. If participants re-partner and optimise, this should also affect the repartition of incomes between spouses. This is precisely what we observe in Figure IV.6.

We present the bivariate-density of individual incomes and household incomes, conditioning on the latter being positive[38]. The x-axis shows participants' labour incomes and the y-axis the households' incomes, each panel presenting the actual treatment status for different family sizes at baseline.

Figure IV.6 shows that among participants, the labour incomes essentially come from their own labour incomes as the 2D-density aligns with the 45° line, contrasting with other comparison groups. Control and never-takers have thicker densities for *bread-earner* households type: positive household incomes with 0 labour incomes from mothers' employment. Those figures show that participants are more likely to be the sole-earner of the household, bunching at the *sweet spot* with far less variability than in the comparison groups. In particular, the figures for parents of 2 show that incomes in the control group go much higher in both dimensions than either the never-takers or treated ones. However, these differences could also partly be due to changes in the share of parents who re-partnered.

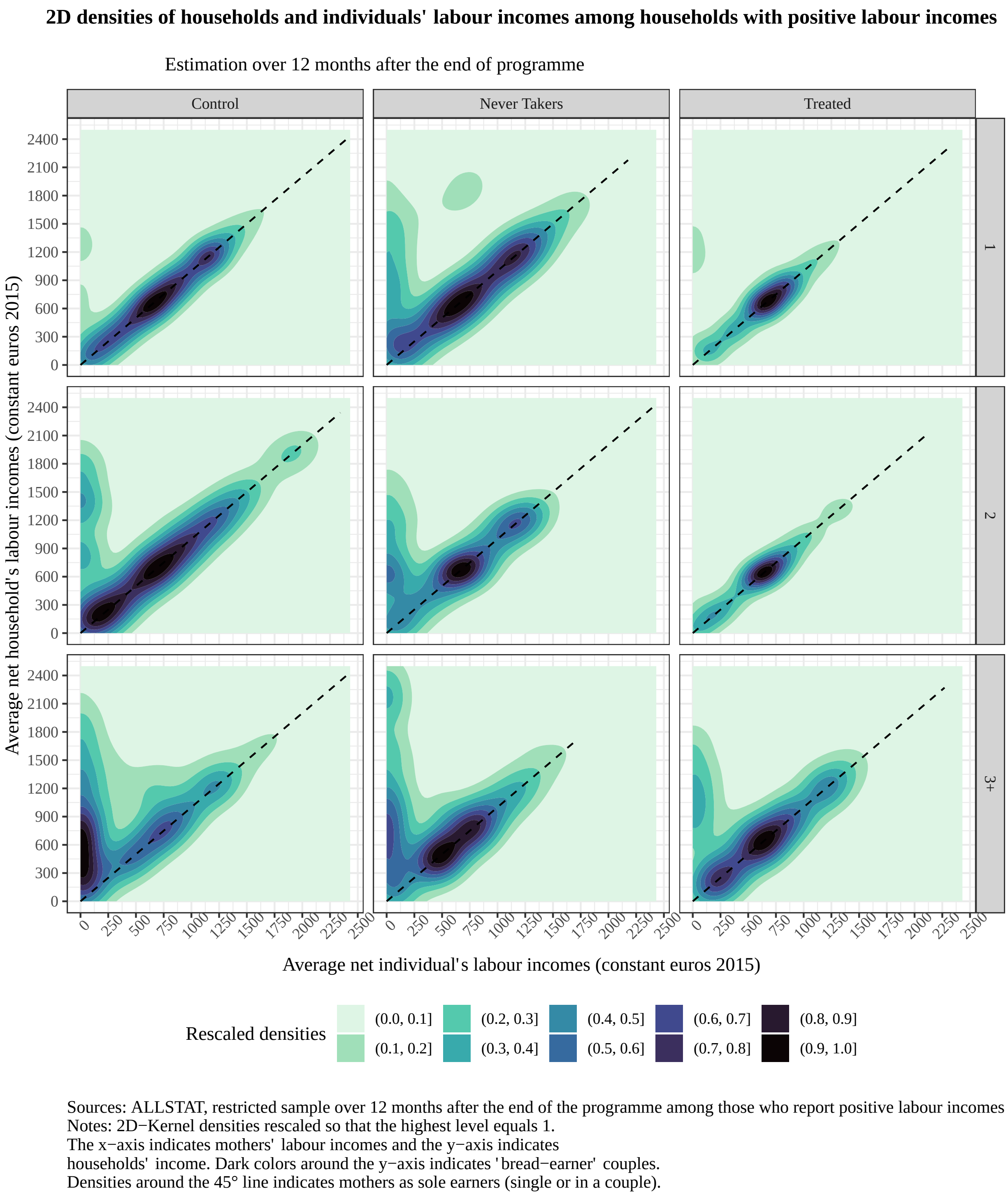
[38] In other words, we use the same sample as in Figure IV.5.

Figure IV.5: Participants bunch at the sweet spot of household labour incomes



Sources: ALLSTAT, restricted sample over 12 months after the end of the programme among those who report positive labour incomes and smaller than 5 000 euros for clarity.
Notes: Kernel density of labour income for those with positive labour incomes. The PA reference line indicates the theoretical amount of in−work benefits and housing benefits received for single parents by number of children and net labour income based on EDIFIS using the 2022 legislation.
Kink points indicate the level of income that minimises the implicit marginal tax rate.

Figure IV.6: 2D density plot of households and individual labour incomes



Small families expands, mothers of two are less pregnants, large families hold their own We now look at the evolution of the family structure variables defined in subsection IV and compare means from 12 months before random assignment to 30 after for the first 4 cohorts. Figure IV.7 estimates the averages using OLS without constant and a dummy for each treatment arm $\times$ relative month, weighting observations with the inverse propensity score of encouragement[39]. We present results for cohabitation, number of dependent children, number of children under 2 and pregnancies, by number of children at baseline.

These figures show at least one important difference in each group and in particular:

- higher cohabitation for parents with one child,
- lower pregnancies for parents of two, and
- higher number of children for those with three or more.

For single parents with three children, the only difference between the encouraged group and the control group is that the average number of children decreases at a slower pace in the former than in the latter.

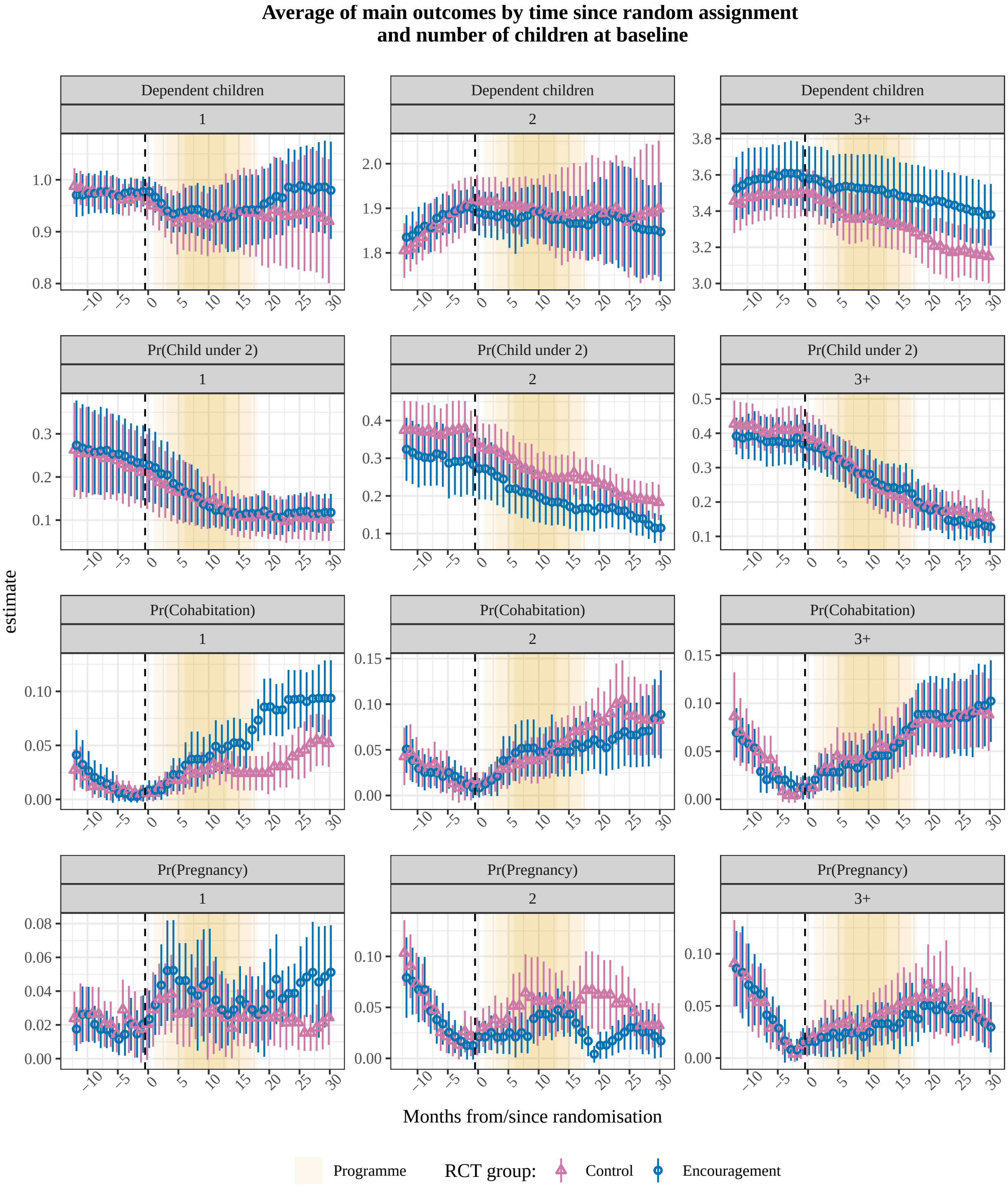
Single parents with one child in the encouragement group are significantly more likely to re-partner, and this occurs right at the end of the training programme. The control group for this type of family remains significantly more single than other family structures. There is no difference between the encouraged and control groups for the number of children under two, but there is an imprecise five percentage point difference in the total number of children. With a difference of about 7 percentage points in cohabitation, this suggests the merging of single-parent families. Later, there is a 2 percentage point difference in pregnancy rates at the end of the timeframe.

Single parents with two children in the encouragement group were less likely to be pregnant during the training period, and the number of pregnant women in this group fell to 0 right after the end of the programme. This period also coincides with a slightly lower share of coupled households, but this difference swiftly fades out. The effect on pregnancy is only slightly starting to show up in the number of children under two at the end of the timeframe.

These descriptive statistics inform on the heterogeneity in the average intention-to-treat effects of the programme on family structure, by number of children at baseline. Importantly, these changes affect the eligibility and amount of transfers received. Figure C.22 in the Appendix also presents the average long difference of the share of families with different sources of incomes, over relative time since random assignment and by number of children. It shows that on average, there are few differences by encouragement group, with two exceptions. First, mothers of three or more children in the encouragement group are more likely to receive family allowances than those in the control group. Second, there is a short lived effect of the programme on the probability of receiving child support, and a symmetric effect on family support allowance for parents of two children.

[39] We estimated a Probit of encouragement on block fixed effect and predicted the assignment probabilities to construct inverse weights. See Heim (2024) for more details.

Figure IV.7: Faster cohabitation for mothers of 1 child, slower reduction of the number of children for mothers of 3 or more



Sources: ALLSTAT 2017−01−01 to 2023−06−01 cohorts 2018 to 2021 from −12 to 30 months from random assignment.
Means and point−wise 95% confidence intervals estimated by OLS regressions on month x group x encouragement dummies without intercept, using cluster−robust standard errors adjusted at the block x cohort level and inverse instrument propensity score weighting.
Rows correspond to different outcomes from separate regressions. Columns display results by number of children at baseline.

V Empirical stategy

We want to estimate how the programme affected single parents' capabilities by observing changes in their income distributions and relate these changes to the economic incentives we identified. In particular, we want to estimate the difference in densities of participants around kink points of the tax rate benefit system with the relevant counterfactual densities. Second, we want to estimate the average treatment effect of the programme on family structure by sub-group of number of children at baseline. We access rich panel data from administrative records of randomly-sampled cohorts assigned using an experimental design. We first start by introducing notations and our main assumptions steaming from this research design. Then, we present our strategy to estimate counterfactual densities, bunching estimator and elasticities in subsection 2. The estimations of dynamic treatment effects on family structure is presented in subsection 3.

1 Notations and definitions

We observe a random sample of individuals $i \in I$ over calendar months $t \in \mathcal{T}$ and we set $t = 1 \equiv$ January, 2017 and $max(\mathcal{T}) =$June, 2023. I is composed of cohorts $c \in \mathcal{T}$ defined by the month before random assignment. We are interested in the effects of the programme aggregating cohorts in relative time and define $m = t - c$, such that $m = 1$ is the month of random assignment. Individuals i in cohort c are defined by a set of attributes including number of children (1,2,3+), registration at the Employment agency (True/False) and number of years receiving RSA (2-5, 5-10, more than 10 years). The cross-product of these characteristics defines blocks $b_{jc} \in B_c \subset B$. B_c is the set of blocks j in cohort c and B contains the $j \times c$ blocks of all cohorts. We denote f_B and F_B the distribution and cumulative distribution of blocks, integrating over b. For estimations, we also denote $b_{ijc} = \mathbb{1}(i \in b_{jc})$ the dummy for block b_{jc} and $\mathbf{B}$ the matrix of block dummies. Number of children at baseline is contained in $\mathbf{B}$ and we denote $\mathbf{B_f} \subset \mathbf{B}$ the set of blocks (and individuals) with families of size f at baseline.

Let $Z = \mathbb{1}(Encouragement)$ and $D = \mathbb{1}(Participant)$ be the random variable for encouragement and participation, and let $\mathbf{Y}$ denote the matrix of outcomes Y^k where k index the outcome of interest. Some outcomes Y^k are *decision* outcomes $i.e.$ they lie in the causal path between D and a final outcome $Y^{k'}$. For instance, taking a job is an outcome possibly affected by the treatment and affecting disposable incomes. In such cases, we let $W \equiv Y^k$ denote this intermediary *decision/outcome*.

Let $f_{Z|B}$ and $F_{Z|B}$ denote the density and cumulative distribution of encouragement across blocks, characterising the instrument propensity score. We denote q_b the encouragement probability in block b with $\mathbb{E}[q_b] = .5$ by design. Due to uneven and small block sizes, there are some with slights variations. To account for that, we run a Probit of Z on $\mathbf{B}$ and use the predicted probability $\hat{q_b}$ as instrument propensity score. Like Heim (2024), we define $\tilde{Z_{ib}} = Z_{ib} - \hat{q_b}$ the centred instrument propensity score. In the framework of Borusyak, Hull, and Jaravel (2022), it follows the idea the distribution of *shocks* $\mathbf{q}$ given $\mathbf{B}$,$\mathbf{q}$, denoted $G(\mathbf{q}|\mathbf{B})$ is known. It then follows from their main results that using $\tilde{Z_{ib}}$ as an instrument for D_i ensures conditional independence and identification of causal effects[40].

Finally, for some random variable R, we define the τ-quantile, r_τ, of R as

$$r_\tau = F_R^{-1}(\tau) := \inf\{r : F_R(r) \geq \tau\},$$

where F_R denotes the cumulative distribution of W.

Our database is made of four random samples of cohorts randomised at time $t = c$ observed $\mathcal{T} - c$ periods after random assignment each. We make the following assumption regarding the properties of our sample:

Hypothesis 3.1 (Random sampling)**.** $\forall c$, the sample $\{Y^k_{ic}, Y^k_{ic+1}, \dots, Y^k_{i\mathcal{T}-c}, \mathbf{B_c}, Z_{ic}, D_{ic}\}_{i=1}^{I_c}$ is independent and identically distributed (iid)

[40] Borusyak, Berkeley, and Hull (2023) propose a more intuitive discussion of the results of Borusyak, Hull, and Jaravel (2022) with numerous illustrations including the link with block random assignment with imperfect compliance.

In words, we assume that the sequence of individual observations in each cohort over the time-frame is randomly sampled from a larger population[41]. Assumption 3.1 allows us to view all potential outcomes as random and imposes no restriction between potential outcomes and treatment allocation, nor does it restrict the time series dependence of the observed random variables. This assumption matters in the justification of standard errors with asymptotics based on the convergence to a population parameter[42]. Together with our identification hypothesis 3.2, this allows to use conventional (cluster) robust standard errors with finite sample corrections to approximate the variance of our estimator.

This assumption may feel at odd with the experimental setting for which design-based estimators of the variance-covariance could improve precision (Alberto Abadie et al. 2020 ; Alberto Abadie et al. 2022). Such estimands use uncertainty in the assignment mechanism and the question is about internal validity: do estimates from the sample reflect causal parameters in this sample? External validity bears on the question of whether causal parameters from the sample correspond to the population parameter. Consistent with our approach of using this experiment as a natural experiment to uncover capabilities and reactions to structural parameters of the tax-benefit system, we think this approach is more relevant and easier to interpret with minimal assumptions on the data generating process.

Potential outcomes and identification For now, we abstract the panel nature of our data and consider identification in the cross-sectional setting at some point after random assignment or over a pre-defined subset of periods. We omit the time subscript to alleviate notation and reintroduce them when it matters. We follow the usual potential outcome notations of Joshua D. Angrist, Imbens, and Rubin (1996) and let $D_{ijc}(z)$, $Y^k(d) = Y^k(D(z)) \equiv Y^k(d,z)$ denote the potential participation and outcomes as function of potential participation and encouragement. Participants constitute the usual "compliers" and their proportion estimates the first stage effect. Because no member of the control group enrolled, there are no "always-takers" of the programme and monotonicity trivially holds. Under SUTVA[43], the observed outcomes reveal potential outcomes and we assume no spillover across groups. Formally, $Y^k(1,1)$ is the revealed potential outcome for treated compliers and $Y^k(0,0)$ the revealed potential outcomes for the control group.

By design, Z and $Y^k(0)$ are independent conditional on attributes summarised in the matrix of blocks. Hypothesis 3.2 summarises our main assumptions:

Hypothesis 3.2 (Identification hypotheses)**.**

1. **One-sided non-compliance:** $Pr(D(0)|Z=0) = 1$
2. **Meaningful first stage:** $Pr(D(1)|Z=1) > 0$
3. **Conditional independence and exclusion:** $Y^k(0) \perp Z|\mathbf{B} = b_{jc} \quad \forall\, b_{jc}$ **where** $Pr(D=1|\mathbf{B}=b_{jc}) > 0$
4. **Common support:** $f_{Z|B}(Z=0|\mathbf{B}=b_{jc}) > 0 \quad \forall\, b_{jc}$ **where** $Pr(D=1|\mathbf{B}=b_{jc}) > 0$

Note that in the usual instrumental variable setting, the independence assumption includes potential participation and potential outcome for the treated. However, Frölich and Melly (2013) showed that this is the only conditional

[41] This assumption reflects closely the actual sampling process although it neglects social workers' assessment and exclusion of about 1/5 ineligible households from the initial random samples. We assume that exclusion is consistent across years and that the remaining experimental sample remains representative of the larger population of eligible single parents.

[42] For instance, consider the estimand δ of the difference in mean between the encouragement sample of size N_1 and control of size N_0. The total variance in the sample is given by the Neymann Variance for randomised experiment

$$V^{\text{total}}\left(N_1, N_0, n_1, n_0\right) = \operatorname{var}\left(\hat{\delta} \mid N_1, N_0\right) = \frac{S_1^2}{N_1} + \frac{S_0^2}{N_0} - \frac{S_\delta^2}{n_0 + n_1}$$

with n_0 and n_1 the unknown size of the population and $S_.$ denotes the population variance of the outcome in each group $(0,1)$ and the variance of the treatment effect δ. For fixed N_0 and N_1, if $n_0, n_1 \to \infty$, the total variance and the sampling variance are equal:

$$\lim_{n_0, n_1, \infty} V^{\text{total}}\left(N_1, N_0, n_1, n_0\right) = \lim_{n_0, n_1, \infty} V^{\text{sampling}}\left(N_1, N_0, n_1, n_0\right) = \frac{S_1^2}{N_1} + \frac{S_0^2}{N_0}.$$

See Negi and Wooldridge (2021) and Alberto Abadie et al. (2020) for details.

[43] Stable unit treatment value assumption

independence required with one sided non-compliance. More precisely, under Hypothesis 3.2, the average treatment effect on the treated is identified by the ratio of the intention-to-treat on the first stage (Frölich and Melly 2013):

$$\begin{aligned} ATT = \mathbb{E}[Y^k(1) - Y^k(0)|D=1] &= \frac{\mathbb{E}[Y^k|Z=1,\mathbf{B}] - \mathbb{E}[Y^k|Z=0,\mathbf{B}]}{\mathbb{E}[D|Z=1,\mathbf{B}] - \mathbb{E}[D|Z=0,\mathbf{B}]} \\ &= \frac{1}{Pr(D=1)} \int \mathbb{E}[Y^k(1) - Y^k(0)|\mathbf{B}, D=1] Pr(D=1|\mathbf{B}) dF_B \end{aligned} \tag{3.4}$$

In words, the effect of the programme on the participants is, in expectation, the ratio of the weighted within-block difference in outcomes between encouragement and control over the within-block weighted difference in participation. The weights correspond to the distribution of blocks, but will typically depend on the estimation method and integration over periods.

Identification of the distribution of missing potential outcomes Under the same set of hypotheses, the average missing potential outcome for the treated is identified and, in fact, so is any measurable function $g(\cdot)$ of that potential outcome, as long as $g(\cdot)$ has finite first moment (Frölich and Melly 2013). Formally:

$$\mathbb{E}[g(Y^k(0))|D=1] = \frac{\int \mathbb{E}[g(Y^k(0))|\mathbf{B}=b, Z=0] dF_B(b) - \mathbb{E}[g(Y^k(0)) \times (1-D)]}{Pr(D=1)} \tag{3.5}$$

This result is particularly interesting in our setting. It means that any function of the *missing* potential outcome for treated compliers is identified. In particular, this equation allows identification of the entire marginal distribution of potential outcomes of treated compliers, had they not been treated.

First, note that under SUTVA and exclusion, the marginal distribution of potential outcomes for treated compliers is directly observed. Indeed, Since $Z = D$, under exclusion, Z is ignorable for compliers given $\mathbf{B}$ and $F_{Y_i|Z=1,D=1,\mathbf{B}} = F_{Y_i(1)|D=1,\mathbf{B}}$. What's missing is the counterfactual distribution $F_{Y_i(0)|D=1,\mathbf{B}}$, which is therefore also identified with this instrumental variable strategy. To alleviate notations, we denote $f_{Y^1|D^1}$ the marginal density of **treated compliers**, $f_{Y^0|D^1}$ the counterfactual density for **untreated compliers**. The marginal density of never-takers is $f_{Y^0|D^0}$.

By setting $g(Y_i(d)) = \mathbb{1}(d \times Y_i \leq y)$ for a constant y and each value of d, we obtain the complier cumulative distribution functions of $Y_i(1)$ and $Y_i(0)$ evaluated at y[44]. For instance, if we want to compute the probability that compliers had no labour income, had they not participated, we can define $g(\cdot) = (1-D_i)\mathbb{1}(Y_i \leq \varepsilon)$ with $\varepsilon \to 0$ and run a TSLS on $(1-D_i)$ instrumented by Z_i, with block fixed effects.

Since we want to look at the bunching around the *sweet spot* of labour incomes, we are more interested in the derivative of g $i.e.$ the density of compliers potential labour income. For that, we can set $g(Y_i(d)) = \frac{1}{h}K(\frac{Y_i - y}{h})$, where $K(\cdot)$ is some kernel function with bandwidth h that shrinks to zero asymptotically.

Abbadie's Kappa and weighted distribution There is a close connexion between this result and that of A. Abadie (2003) in the case with two-sided non-compliance. In this paper, Abadie shows that there are weighting representations of the LATE based on quantities he calls κ. Using these weights, one can identify the potential outcome distribution and derivatives for treated and untreated compliers. Alberto Abadie, Angrist, and Imbens (2002) further use these weights to define instrumental quantile treatment effects and implement this method on data from the JTPA, a randomised welfare-to-work programme in the US. We use the following Theorem as our main identification framework:

[44] A very intuitive presentation of this result can be found in the chapter on method for school effectiveness by J. Angrist, Hull, and Walters (2023). Heim (2024) uses this method to estimate the distribution of disposable income over the year after the programme ended and found no differences.

Theorem 3.1 (Weighted representation of expected functions of potential outcomes)**.** *Under Hypothesis 3.2, Frölich and Melly (2013) show that with one-sided non compliance:*

$$\begin{aligned}\mathbb{E}[g((Y(0)|D=1)] = & \frac{1}{Pr(D=1)}\mathbb{E}\Big[g(Y)\times(1-D)\frac{Pr(Z=1|\mathbf{B})-Z}{Pr(Z=1|\mathbf{B})\,(1-Pr(Z=1|\mathbf{B}))}\Big] \\ = & \frac{1}{Pr(D=1)}\mathbb{E}\Big[g(Y)\times(1-D)\frac{-\tilde{Z}}{q_b\,(1-q_b)}\Big]\end{aligned} \tag{3.6}$$

$$\begin{aligned}\mathbb{E}[g((Y(1)|D=1)] = & \frac{1}{Pr(D=1)}\mathbb{E}\Big[g(Y)\times(D)\frac{Z-Pr(Z=1|\mathbf{B})}{Pr(Z=1|\mathbf{B})\,(1-Pr(Z=1|\mathbf{B}))}\Big] \\ = & \frac{1}{Pr(D=1)}\mathbb{E}\Big[g(Y)\times(D)\frac{\tilde{Z}}{q_b\,(1-q_b)}\Big]\end{aligned} \tag{3.7}$$

The proof of Theorem 3.1 is given in Frölich and Melly (2013) and we simply adapt notations to our setting. The second lines of each equation simply replace $P(Z=1|\mathbf{B})$ by the instrument propensity score q_b previously defined. We use this theorem to estimate counterfactual densities of participants. Other recent work revisit the use of κ weights to estimate the LATE (Słoczyński, Uysal, and Wooldridge 2022a). In this paper, attention is drawn to the different estimands for the weights which may be negative and not sum to unity. Then again, one-sided non-compliance have better properties than two-sided non-compliance and in particular, non-compliance ensures positive weights (Proposition 3.3 of Słoczyński, Uysal, and Wooldridge (2022a)). In our application, we simply *plug* estimations of the instrument propensity score q_b to compute these weights (See Subsection 2 below).

Comments: what two-stage least square estimates Consider the following simple TSLS system at any period after random assignment:

$$\begin{cases} Y_{ib} = & \mathbf{B}'\beta_b + \delta D_{ib} + \mu_{ib} \\ D_{ib} = & \mathbf{B}'\alpha_b + \pi\tilde{Z}_{ib} + \epsilon_{ib}\end{cases} \tag{3.8}$$

In this system, block x cohort instrument themselves in the second equation while participation is instrumented by the demeaned instrument. Because blocks are discrete and **B** contains indicators for each possible realisation, the specification is saturated in B. This is the typical TSLS of J. D. Angrist and Imbens (1995) but with centred instrument. As the recent work of Blandhol et al. (2022) shows, TSLS retrieves a LATE interpretation only in saturated specifications, such that the projection matrix of the first stage fits the conditional expectation. Note that the original result on TSLS and the LATE use a fully saturated model in the first stage $i.e.$ interacting the instrument with each set of dummy covariates. This amounts to integrating the covariate-specific LATE over the distribution of covariates with treatment-variance weighting. This "saturate and weight" specification is like nonparametric conditioning but uses only a single treatment variable (Blandhol et al. 2022; J. Angrist and Kolesár 2022). The problem with this specification is that it has many excluded instruments and is more sensitive to both small sample and many-instruments bias. However, Borusyak, Hull, and Jaravel (2022) show that centring on the propensity score recovers the same parameter, while also making explicit the modelling assumption of the first stage to estimate the propensity score q_b. They show that equation (3.8) identifies weighted averages of conditional-on-block IV coefficients[45]:

$$\begin{aligned}\delta = & \mathbb{E}\Big[\frac{\sigma_Z^2(\mathbf{B})\pi_b}{\mathbb{E}[\sigma_Z^2(\mathbf{B})\pi_b]}\delta_b\Big] \\ = & \int_b \frac{q_b(1-q_b)\pi_b}{\int_b q_b(1-q_b)\pi_b dF_B}\times\delta_b\,dF_B\end{aligned}$$

[45] See the chapter by J. Angrist, Hull, and Walters (2023) with a very clear presentation and application to school choice and the recent review by Borusyak, Berkeley, and Hull (2023) for more intuition on the theoretical results and review of other applications.

An important issue is the weighting of this TSLS that is proportional to the conditional variance of the instrument and the conditional first stage π_b. When propensity scores are constant (which is asymptotically our case), $\sigma_Z^2(\mathbf{B}) = \mathbb{E}[\sigma_Z^2(\mathbf{B})]$ and the β_b are weighted only by the conditional complier shares, yielding the unconditional LATE[46]. Note that from an identification perspective, block fixed effects are then unnecessary. We include them to improve precision in the second stage.

2 Estimating distributional effects

To assess the optimisation behaviours of treated compliers, we want to estimate their distribution of potential incomes had they not been treated and compare the bunching around the kinks of the French tax benefit. For that, we define a window of observation with balanced number of months across cohorts and consider the density of incomes over that period. To gain intuition and link our approach to the model presented in Subsection 2 of our theoretical framework, we first present and estimate typical bunching estimators *à la* Saez, Slemrod, and Giertz (2012).

The usual bunching estimator Recall that the objective is to estimate the excess and missing mass of observations around the kink points of the budget set to measure the sensitivity of pre-tax income to variation of tax rate, with varying knowledge of the tax-benefit system between participants and non-participants.

In the bunching literature, one usually does not have comparison groups, and the counterfactual density needs to be estimated from the same observations (or comparing densities around a reform), usually using a parametric polynomial regression of the density.

We follow standard practices[47] and define bins H indexed over j with fixed width h set as 5% minimum wage around the bunching point X^*. The latter is at 60% minimum wage but we also run models at other kinks. We then estimate the following model:

$$C_j = \sum_{\upsilon=0}^{p} \beta_\upsilon \left(H_j\right)^\upsilon + \sum_{x=H_L}^{H_U} \gamma_x \mathbb{1}\left[H_j = x\right] + \sum_{r\in R} \rho_r \mathbb{1}\left[\frac{H_j}{r} \in \mathbb{N}\right] + \sum_{k\in K} \theta_k \mathbb{1}\left[H_j \in K \wedge H_j \notin [H_L, H_U]\right] + \mu_j \tag{3.9}$$

C_j is the observation count in bin j, p is the order of polynomial used to fit the counts, and H_L and H_U stand for the lower and upper bins that define the bunching region. We define these parameters by visually inspecting the bunching region to account for imprecise bunching. In practice, we use a degree-4 polynomial.

This specification corrects for round numbers R (defined by $r \in \mathbb{N}$), and for other fixed effects in a set K that feature a bunching mass in the estimation bandwidth outside the bunching range $z \in [H_L, H_U]$ but that are not associated with X^* (through the ρ and θ coefficient vectors respectively). In particular, weekdays are susceptible to create bunching mass every 20% of the minimum wage and there may be other bunching mass due to other kinks/notch or reference point. Not controlling for round number bunching can significantly bias the bunching estimate upwards since $X^* = 60\%$ is also a round number. This is because some of the observed bunching will be driven by factors unrelated to the change in incentives driven by the discrete change in the constraint that we want to attribute the bunching to. Similarly, not controlling for other bunching masses can exert a downward bias in the bunching estimate at X^* by biasing the counterfactual estimate upwards.

Given this estimation strategy, the predicted counterfactual density in the absence of the kink is given by:

$$\hat{B}_0 = \sum_{j=H_L}^{H_U} \left(C_j - \hat{C}_j\right)$$

To be able to compare bunching masses across different kinks or groups featuring varying heights of counterfactuals, we use a normalisation where we divide the total excess mass with the height of the counterfactual at X^*. The

[46] A very clear note on this transformation can be found on the web page of Peter Hull: https://about.peterhull.net/metrix

[47] We use the `Bunchit` package that directly implements the estimates *à là* R. Chetty et al. (2011) and Henrik J. Kleven and Waseem (2013).

normalised mass is:

$$\hat{B} = \frac{\hat{B}_0}{\hat{C}_0}$$

In the model presented in Section 2, this parameter corresponds to the following approximation:

$$B = \int_{X^*}^{X^*+\Delta X^*} h_0(z)dz \approx C_0(X^*)\Delta X^*$$

From there, we can estimate the elasticities using the value of the tax rate around the kink and pluging the density estimates in equation (3.2) presented in the theoretical framework. This approximation is only valid for small changes in the marginal tax rate so instead, we assume iso-elastic utility functions and retrieve the parametric observed elasticity with this simple expression[48]:

$$e = -\frac{\ln\left(1 + \frac{\hat{B}}{C_0(X^*)X^*}\right)}{\ln\left(\frac{1-\tau_0}{1-\tau_1}\right)} \tag{3.10}$$

We estimate such models separately for the treated and untreated participants and by number of children over the post-treatment period and see descriptively how these distributions diverge and what the implied elasticities are. We use bootstrap to compute standard errors for the bunching mass and elasticities.

These models use parametric assumption and data from the same group to estimate the counterfactual density. Among participants, the bunching mass corresponds to optimising families with ability affected by the programme, higher knowledge and possibly lower adjustment costs. Conversely, non-participants contain twice the share of never-takers population and the untreated compliers. Using the identification results of Theorem 3.1, we can instead use the counterfactual density of untreated compliers to estimate these elasticites. We now define our preferred strategy to retrieve the counterfactual densities.

Local regression distribution Our main results are based on the so called *"local regression distribution estimator"* proposed by Matias D. Cattaneo, Jansson, and Ma (2021) and implemented with the R package `lpdensity`. A very interesting feature of this estimator is that it is semi-parametric and data driven. It uses local polynomial regressions with MSE-optimal bandwidth. The authors also define point-wise and uniform confidence intervals using bootstrap the package implements. Another important advantage of this estimator is that it does not require preliminary smoothing of the data and hence avoids preliminary tuning parameter choices.

An important condition for estimation is that $F(\cdot)$ is suitably smooth near y. The function solves for parameters $\theta(\mathbf{y})$ using the local regression estimator:

$$\hat{\theta}(\mathbf{y}) = \underset{\theta}{\text{argmin}} \sum_{i=1}^{n} W_i \left(\hat{F}_i - R_i'\theta\right)^2, \tag{3.11}$$

where $W_i = \frac{1}{h} K\left(\frac{y_i - y}{h}\right)$ for some kernel $K(\cdot)$ and some bandwidth $h, R_i = R(y_i - y)$, and

$$\hat{F}_i = \frac{1}{n} \sum_{j=1}^{n} \mathbb{1}(y_j \leq y_i)$$

[48] To get to this result, denote the ability level of the marginal buncher by a^*, his optimal earnings are $x^* + \Delta x^* = n^*(1-t_0)^e$ under the linear schedule and $x^* = a^*(1-t_1)^e$ under the kinked schedule. Combining these gives:

$$e = -\frac{\ln\left(1 + \frac{\Delta X^*}{X^*}\right)}{\ln\left(\frac{1-t_0}{1-t_1}\right)}$$

is the empirical distribution function evaluated at y_i, with $\mathbb{1}(\cdot)$ denoting the indicator function. The generic formulation (3.11) is particularly interesting when $F(\cdot)$ is sufficiently smooth, in which case, the Taylor expansion gives:

$$F(y) \approx F(x) + f(x)(y-x) + \cdots + f^{(p-1)}(x)\frac{1}{p!}(x-x)^p \quad \text{for } y \approx x, \tag{3.12}$$

and $f^{(s)}(x) = \frac{\mathrm{d}^s f(y)}{\mathrm{d}y^s}\Big|_{y=x}$ are higher-order density derivatives. The approximation (3.12) is of the form (3.11) with $R(u) = (1, u, \ldots, u^p/p!)'$, and hence $\theta(x) = \left(F(x), f(x), \ldots, f^{(p-1)}(x)\right)'$ (Matias D. Cattaneo, Jansson, and Ma 2021).

More importantly, their estimator also applies to the class of generic weighted distribution parameters

$$H(\mathbf{x}) = \mathbb{E}[w_i \mathbb{1}(x_i \leq \mathbf{x})]$$

where weights w_i can be used to retrieve the distribution of treated and untreated compliers using the results of Theorem 3.1. The weighted cumulative distribution function is defined similarly:

$$\hat{F}_{wi}(y) = \frac{1}{n}\sum_{j=1}^{n} \mathbb{1}(w_i y_j \leq y)$$

To estimate the distribution of potential income for treated compliers $i.e.$ $f_{Y^1|D^1}(y)$, or their unobservable potential outcomes $f_{Y(0)|D_1}(y)$, the distribution of compliers had they not been treated (what we call *untreated-compliers*), Theorem 3.1 shows that choosing $g(\cdot) = \hat{F}_{wi}(y)$ and using the κ-weighting schemes retrieve the expected density or cumulative CDF of potential outcomes. With a known design, we can simply plug the value of the instrument propensity scores $\hat{q}$ or centred instrument $\tilde{Z}_i$ in the formula. In their empirical application, Matias D. Cattaneo, Jansson, and Ma (2021) do exactly that with data of the JTPA, a very similar setting. Formally:

$$\begin{cases} \kappa_0 = \frac{1}{1-D_i}\frac{1-Z_i-(1-\hat{q}_b)}{\hat{q}_b(1-\hat{q}_b)} \\ \kappa_1 = \frac{1}{D_i}\frac{Z_i-\hat{q}_b}{\hat{q}_b(1-\hat{q}_b)} \end{cases}$$

Joint visualisation of intensive and extensive margin reaction Finally, this method also has the advantage of allowing to rescale the density over a subset of the support of the outcome. This is interesting because we need the distribution to be bounded away from 0 to be estimated with this estimator, and we can estimate the mass point at 0 for the counterfactual distribution. For that, we need estimates of $\mathbb{E}[Y(0) = 0|D = 1]$ and $\mathbb{E}[Y(1) = 0|D = 1]$. The latter can be immediately computed while for the former, we can also use Theorem 3.1 and set $g(\cdot) = (1-D_i)Y_i$. In practice we use a simple TSLS regression of $(1-D_i)Y_i$ on $(1-D_i)$ instrumented by $\tilde{Z}_i$ and block fixed effects. By scaling the weighted density over the support $1-\mathbb{E}[Y(\cdot) = 0|D = 1]$ for each density, we can jointly observe the excess and missing densities over the rescaled support.

Remarks Our data and experimental design allow us to estimate potential outcomes densities corresponding to our treatment: the Reliance programme. It allows to retrieve the distribution of compliers had they not been treated and therefore, provides new measurements of the selection bias strongly emphasised in Heim (2024). For that, we can compare the density of never-takers $f_{Y^0|D^0}(y)$ directly observed in the encouraged group with the counterfactual density of untreated compliers $f_{Y(0)|D_1}(y)$. We can therefore see if without the treatment, compliers' income distribution differs from never-takers.

By comparing the density mass around the kink point for treated and untreated compliers and never-takers, we can measure how the programme affected optimisation behaviours and use these estimates as bunching estimators. The general idea is that without strong simplification such as those imposed in the model of Subsection 2 of our theoretical framework, the observed bunching may confound the financial incentive effect with the reference point of part-time minimum jobs as well as optimisation friction and knowledge costs. If we are willing to make additional assumptions as to which parameter was affected by the programme and which remains constant in comparison of never-takers and untreated compliers, we could recover more structural parameters.

Finally, the comparison of the distribution of potential outcomes for individual and household income also shed light on the evolution of within-household income sources and the role of partners' income. We now briefly describe our estimation strategy for the outcomes on family structure.

3 Treatment effects on the treated

To assess the dynamic treatment effect of the programme on cohabitation and the number of children, we estimate the $ATTs$ defined in equation (3.4) for some periods $S(m)$. For that, our approach simply consists in stacking the TSLS systems of equations (3.8) for each period $S_s(m)$, with $s \in S$ by estimating the model over $S_s(m) \quad s.t.$ $m \in \{-12, ...\mathcal{T} - max(c)\}$ interacting all right hand side elements with a set of S dummies for $S(m)$. Formally, we estimate:

$$\begin{cases} Y_{ibm} = & \sum_s \beta_{bs} \mathbf{B}' S_s(m) + \sum_s \delta_s D_{ib} S_s(m) + \mu_{ibm} \\ \sum_s D_{ib} S_s(m) = & \sum_s \alpha_{bs} \mathbf{B}' S_s(m) + \sum_s \pi_s \tilde{Z}_{ib} S_s(m) + \epsilon_{ibm} \end{cases} \tag{3.13}$$

In words, there are S first stages, one for each period s where participation D is instrumented by the corresponding interaction between the centred instrument and a period dummy. Fixed effects interacted with period dummies instrument themselves in the second stage and allow for block $\times$ period specific differences *between* blocks. As noted by Blandhol et al. (2022) "*nonparametric TSLS specifications that restrict attention to the population with $X = x$ and use a first stage that is saturated in Z will be both rich (trivially) and monotonicity-correct. Each value of x produces a different estimand $\beta_{TSLS}(x)$, each of which is weakly causal. Any positively weighted sum of $\beta_{TSLS}(x)$ across $x \in X$ will be weakly causal*".

The collection of $\hat{\delta}_s$ estimate the $ATTs$. If $m < 1$, the estimations are $placebo$ estimates and should be close to zero. We restrict such test above -12 as we can only observe the first cohort that many leads before random assignment. Similarly, we only use observations up to the maximum leads for the last cohort to have a balanced composition of groups over the periods.

The equivalent reduced form equation is:

$$Y_{ibm} = \sum_s \lambda_{bs} \mathbf{B}' S_s(m) + \sum_s \rho_s Z_{ib} S_s(m) + \upsilon_{ibm} \tag{3.14}$$

In which case, the collection of estimated parameters $\hat{\rho}_s$ are intention-to-treat effects at period p.

We use cluster-robust standard errors adjusted at the block level to account for possible heterogeneity within block following C. de Chaisemartin and Ramirez-Cuellar (2022). We also account for simultaneous inference by constructing a uniform 95% confidence interval for all $\hat{\delta}_s$ (resp. $\hat{\rho}_s$) using the Holm-Bonferroni correction (Hothorn, Bretz, and Westfall 2008).

Instrumented difference-in-differences The previous model is fully saturated and computes one coefficient per period $S_s(m)$. Alternatively, we could omit the coefficient corresponding to the period before random assignment and estimate an instrumented $event - study$ design. This model is just identified, while the interaction of block fixed effects with period dummies ensures clean comparisons. It assumes parallel trend in both the equivalent reduced form (ensured by random assignment) and the first stage. In our setting, once participants are assigned treatment, there is no switch-back, ensuring a constant share of treated units at every period. Under such settings, Clément De Chaisemartin and D'Haultfœuille (2017) show that the TSLS correspond to the LATE.

Heterogeneous treatment effects We often want to identify the effect of the programme on participants with different baseline characteristics; number of children in particular.

Because this variable is part of those defining blocks, the conditional treatment effects on participants is simply identified over the subset of blocks whose participants have the characteristic of interest.

To simultaneously estimate the models by groups of family size $\times$ period, we estimate the following equations by TSLS:

$$\begin{cases} Y_{ibm} = & \sum_s \beta_{bs} S_s(m)\mathbf{B}' + \sum_s \sum_f \delta_{sf} \mathbf{B_f}' D_{ib} S_s(m) + \mu_{ibm} \\ \mathbf{B_f}' D_{ib} S_s(m) = & \sum_s \alpha_{bs} \mathbf{B}' S_s(m) + \sum_s \sum_f \pi_{sb} \mathbf{B_f}' \tilde{Z}_{ib} S_s(m) + \epsilon_{ibm} \end{cases} \tag{3.15}$$

On the right hand sides, we still have block fixed effect interacted with each period. Number of children are part of these blocks. Then, we have a double interaction between treatment, number of children in the matrix $\mathbf{B^f}$ and either treatment dummies in the second stage, or the demeaned instrument in the first. This estimates one average treatment effect on the treated per family structure at period $S_s(m)$. To correct for simultaneous inference, we consider the uniform 95% confidence interval over periods, but within group of family size.

VI Optimisation of income sources and number of shares

As announced in the previous section, we first start by presenting bunching estimates comparing the distribution of individual labour incomes reported over the 12 months from 18 to 30 months after random assignment. Then, we present our main results comparing counterfactual densities using weighted regression distributions. While we always report results for the entire sample, most of the analyses focus on heterogeneous responses by number of children at baseline.

1 Bunching estimates

In order to estimate observable elasticities and understand the magnitude of the change in implicit tax rate single parents face, we need tax-rate parameters. For that, we use the EDIFIS model and use rough average tax rate for calibration. Elasticities using bunching estimates rely on strong assumptions and we do not interpret them as structural parameters. However, they offer a simple metric for comparisons across groups. Table VI.1 presents the average marginal tax rate over (large) bins of household labour incomes by household size, relationship status and additionally, we include simulation when the partner earns either nothing (RSA) or a full-time minimum wage. The continuous implicit marginal tax rate is presented in Figure A.19 in the Appendix.

Aggregate implicit tax rates This table shows the sharp variations around the identified kink points (60 and 100% full-time minimum wage) and differences across family structures for different combinations of incomes. In particular, single parents with one or two children are more heavily taxed on almost every bins of the (low) earning distribution. Furthermore, their implicit tax rate is very high except the *sweet spot* hole. This comes from the fact that the PA has individual bonuses starting at .5 minimum wage, sharply reducing the implicit tax rate for dual-earners households from 50% to 100% MW. Above minimum wage, the IMTR increases for everyone, but is much higher for single parents or single-earner couples.
Single parents with three children have a rather decreasing marginal tax rate except the notch at 60% minimum wage and above where they receive no more PA. It is also striking to see that on average, the implicit marginal tax rate at the *sweet* spot decreases with number of children but is the same with or without a partner if he does not work. However, because of the individual bonus, it is even less taxing for the household to remain at the sweet spot with a partner with a full-time minimum wage.

Table VI.1: Average implicit marginal tax rate by family structure and earnings

N children	Relationship	Partner's income	Earned income bin							
			[0;40[	[40;49[	[50;60[	[60;80[	[80;100[	[100;110[	[110;120[	[120;150[
1	Single	No partner	39	66	24	46	46	74	55	39
	Couple	RSA	39	39	24	62	46	64	39	39
		SMIC	45	39	13	17	22	39	39	39
2	Single	No partner	39	71	19	42	43	70	67	49
	Couple	RSA	39	39	19	61	43	70	67	49
		SMIC	61	39	13	13	13	39	39	56
3	Single	No partner	49	39	13	38	31	26	24	26
	Couple	RSA	39	39	13	61	38	66	63	88
		SMIC	64	63	40	51	13	39	39	39
	Single	Average	42	59	19	42	40	57	49	38

Source: EDIFIS, 2022 legislation. Simulation based on children between 3 and 14, single parent receiving ASF and no other income. Mean implicit marginal tax rates in 2022 by family structure. All rates have been rounded to the unit. The last row averages rates for single mothers of 1, 2 or three children.

Parametric bunching estimates We present our estimates of the bunching densities of household pre-tax incomes around the 60% kink point for participants and non-participants in Figure VI.8. We also report the estimate of bunching at the full-time minimum wage that we comment after. All models use bins of 5% minimum wage and a 4-degree polynomial fit with dummies for round numbers every 20% minimum wage. The predicted counterfactuals are displayed in red. For the bunching around 60% minimum wage, we enlarge the bunching region to 2 bins on the left and 1 on the right, allowing imprecise bunching. These choices were made observing the plots, and means that we compute marginal bunchers from the region between 50 and 65 % minimum wage.

In the top panel, we observe the number of participants reporting incomes in each bin over the 12 months period, with massive bunching of observation in the *sweet spot*. Compared with the polynomial fit, there is a missing density before, another excessive density around 70-75% minimum wage then a rapidly fading mass of reported labour incomes. Conversely, the density of participants exhibits much lower and diffuse bunching mass around the *sweet spot*, but another bunching mass at full-time minimum wage. This is why we also present a model for the bunching mass among non-participants around the full-time minimum wage. For this model, we add another fixed effect at 60% minimum wage to capture the first density mass, 1 bins left and 2 bins on the right.

These estimates are similar to the kernel density presented in Figure IV.5 in Section 2, but use a parametric model to predict the counterfactual from the density of observations used in each model, excluding the bunching region.

Interpretation and important caveats It is striking how much labour incomes are missing above 75% minimum wage among participants. Compared with non-participants, most incomes above that threshold have been shifted to the bunching region. For non-participants however, the very different distribution needs to be discussed.

Before going into more economic interpretations, we should recall important caveats regarding these estimations and their underlying assumptions.

An important aspect of bunching estimate is that it requires structure on preference, ability or other parameters to be informative about the underlying structural parameters. This important result has been showed by Blomquist and Newey (2017). Without *a priori* knowledge of the elasticity, the size of the interval and the quality of the approximation are both unknown. The elasticity is sensitive to the shape of the underlying distribution and tuning the polynomial order and bunching window can lead to substantial variations in the estimations (Bertanha et al. 2023).

There are therefore three possible explanations for the shape of the distribution of income among non-participants. 1) higher friction costs ψ_i, 2) different knowledge and perception of the tax schedule θ_i and 3) differences in ability a_i.

Regarding point 1), we showed that the full-time minimum wage is the most heavily taxed for single parents which should create bunching below. However, the minimum wage sets the lowest wage rate and earnings can only be lower through a part-time job. This may create sparser sets of choices, over number of worked days for instance[49]. In the literature, most empirical estimates use discrete choice models both as a facilitating assumption for estimation but also because it better reflects the institutional setting in the labour market (Briard 2020).

Regarding point 2), part-time and full-time minimum wages are important reference points and the bunching mass may also reflect the institutional setting of the labour market, and how people and society value mothers at work (Maurin and Moschion 2009; Cassar and Meier 2018; Fleche, Lepinteur, and Powdthavee 2018a; Cavapozzi, Francesconi, and Nicoletti 2021; Schmidt et al. 2023). However, it may also be that single mothers do not know about tax rate variations and make their decision as if the tax rate was linear. Liebman and Zeckhauser (2004) use the term "*schmedule*" and the action of "*schmeduling*" to designate economic agents' inaccurate perception of tax-schedule, causing two typical schmeduling behaviours: "*ironing*", i.e. making decisions based on the average price at the point where they consume, or "*spotlighting*", when agents identify and respond to immediate or local prices, and ignore the full schedule, even though future prices will be affected by current consumption. For instance, Rees-Jones and Taubinsky (2020) analyse these perceptions of tax schedule in the US and find that 43% of the population *irons*. Conversely, participants had training sessions and simulations and may be less likely to *schmedule* and be closer to frictionless optimisation. We discuss point 3) in the next paragraph.

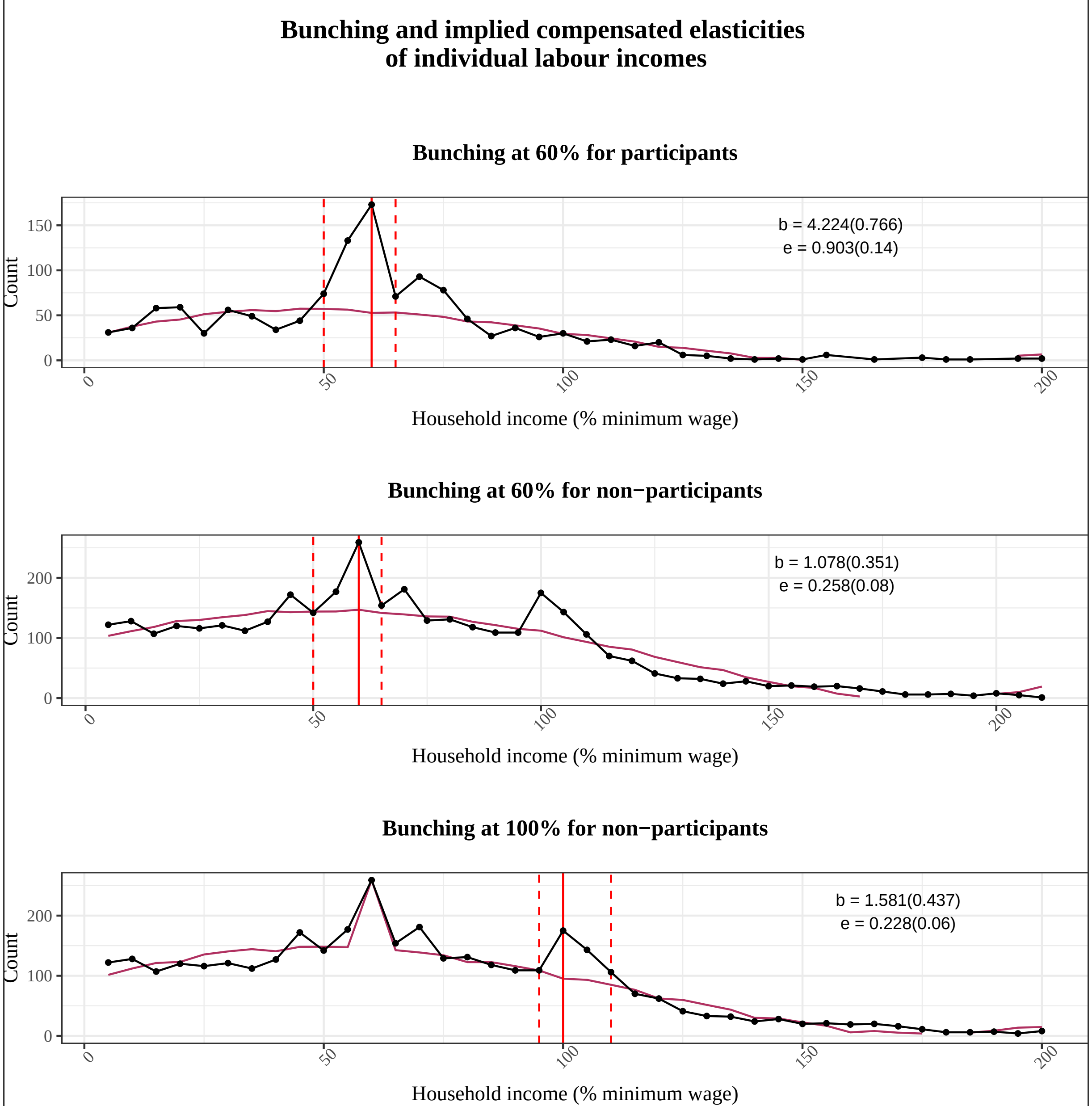
Implied Elasticities of labour incomes An important difference between participants and non-participants is the composition of their groups in terms of latent compliance type. The distribution of non-participants includes untreated compliers and twice the group of never-takers. These two populations may differ in terms of potential outcomes *i.e.* $f_{Y^0|D^1} \neq f_{Y^0|D^0}$. The distribution we observe and the counterfactual are a mixture of both groups, and is therefore not an appropriate counterfactual for treated compliers. However, this mixture of types may explain the second bunching at the minimum wage if we think never-takers and untreated compliers differ in term of *ability* a_i, as defined in the theoretical model. In this case, the prediction of this model is that individuals with higher ability chose a level of labour income higher than the kink region and are therefore insensitive to it. However, for those who would chose an optimal level of labour income slightly above the minimum wage in the absence of a kink, these can be the bunchers we observe[50].

If we are willing to make this different ability assumption and also that their utility is iso-elastic, then this bunching mass identifies the elasticity of individual income to the implicit tax rate for this particular group. We use the tax rates presented in the last row of Table VI.1 to compute the observed elasticities in both groups around the kink. For non-participants around full-time minimum wage, the observed elasticity is 0.23, which is rather low, and consistent with other estimates in the literature (Briard 2020; Bargain et al. 2014). Interestingly, we observe roughly the same elasticity when we use bunching at the 60% minimum wage kink (0.26) for this group.

[49] This is why we use dummies for round numbers every 20% of the minimum wage in the estimation of counterfactual densities

[50] This would be consistent with the result in Heim (2024) showing that compliers are more likely to among the lowest education group.

Figure VI.8: Large changes in the distribution of individual labour incomes



Sources: ALLSTAT 2017−01−01 to 2023−06−01 . Observations from 18 to 30 months after random assignment.
The database is collapsed into frequencies of individual labour incomes falling into bins of 5% minimum wage.
The statistical units are individual x months with positive labour incomes.
The top panel is from sub−sample of participants, the bottom panels are from the control group and never−takers.
Counterfactual densities estimated using polynomial of degree 4 , with round number dummies for every 20% SMIC corresponding to round number of days worked.
The bunching regions are delimited by the dashed red lines and we use the average IMTR for single parents with 1,2,3 chidren presented in the table at the begining of the section.
Elasticities use parametric restriction and standard errors are computed by bootstrap using 499 replications.

The elasticity that we compute with the same structural hypotheses for participants is however much higher. For the latter, the marginal buncher is located at 81% of the minimum wage, and the implied elasticity is 0.9, 3.5 times higher than the elasticity for non-participants at this kink point. The marginal buncher for the latter is located at 65% of the minimum wage.

Figure D.24 in the Appendix compares the bunching estimates by encouragement status and an estimate for never-takers, which reveal interesting results. Never-takers exhibit a rather sharp bunching at the kink point on an otherwise very rather uniform distribution until 110% minimum wage, where the density sharply decreases. Compared with the density in the control group – in witch they represent 61% of households –, these two figures suggest that untreated compliers would have been the one to bunch on the full-time minimum wage. The weight of never-takers in the control group explain most of the bunching mass at the 60% kink point observed in the control group. However, never-takers do not bunch at the full-time minimum wage.

We complete this picture by estimating the model separately by number of children at baseline and report the results in Figure D.25 in the Appendix. There are three important results we should note.

1) The entire excess mass among non-participants around full-time minimum wage comes from single parents who had one child at the time of random assignment.
2) Non participants with two children bunch sharply at part-time minimum wage.
3) Among participants, the bunching concerns every groups although there are far fewer parents of three with positive labour incomes (in both groups).

These first results confirm the sharp differences observed in the descriptive statistics and provide first estimates of the labour elasticities between participants and non-participants. Without the programme, single parents on welfare seem pretty un-reactive to incentives of the tax-benefit system. However, the counterfactual densities are estimated using parametric assumptions within these groups so their interpretations rely on strong hypotheses. We now present our semi-parametric approach to estimate the counterfactual densities of untreated compliers.

2 The distribution of potential outcomes of treated and untreated compliers

To estimate the densities for treated and untreated compliers, we use the new method of weighted and rescaled regression distribution of Matias D. Cattaneo, Jansson, and Ma (2021) presented in section V.

To start, we measure the bunching mass at 0 income for treated and untreated compliers by TSLS and report these estimates in Table VI.2. Each column is a separate outcome and we estimate jointly the potential outcomes by number of children using stacked regressions. First, the more children they have, the less likely they are to report labour incomes of their own. Second, as already shown by Heim (2024), the programme has a negative effect at the extensive margin for single parents of one child: 76% of participants reported no individual labour income over the year, but they would have been 61% had they not participated. These significant differences are also seen on household labour and individual incomes. However, we see no significant difference in potential null income for parents with more children.

On average, about half of single parents with one or two child(ren) reported some individual incomes over the year, and between 2/3 and 3/4 did not report any labour income. We use these estimates to rescale the densities estimated over positive incomes and use the weighted local distribution regression using the estimator of Matias D. Cattaneo, Jansson, and Ma (2021) presented in Section V. We report estimates for individual labour incomes in the body of the paper and provide estimations for other outcomes in Appendix E.

Table VI.2: Average potential probability of 0 income over 30 months

	Labour income		**Pre-tax income**	
	Individual	Household	Individual	Household
1(Y(0)=0):Children=1	0.614	0.591	0.446	0.449
	[0.526, 0.702]	[0.496, 0.685]	[0.330, 0.561]	[0.320, 0.577]
1(Y(0)=0):Children=2	0.739	0.711	0.511	0.447
	[0.622, 0.856]	[0.592, 0.829]	[0.261, 0.760]	[0.223, 0.671]
1(Y(0)=0):Children=3+	0.896	0.728	0.701	0.556
	[0.775, 1.017]	[0.604, 0.852]	[0.529, 0.874]	[0.400, 0.713]
R2	0.497	0.407	0.293	0.253
R2 Adj.	0.496	0.407	0.292	0.252
1(Y(1)=0):Children=1	0.759	0.714	0.569	0.537
	[0.706, 0.812]	[0.661, 0.767]	[0.486, 0.652]	[0.454, 0.620]
1(Y(1)=0):Children=2	0.721	0.676	0.488	0.450
	[0.661, 0.780]	[0.610, 0.742]	[0.422, 0.554]	[0.387, 0.513]
1(Y(1)=0):Children=3+	0.844	0.777	0.606	0.537
	[0.799, 0.889]	[0.723, 0.830]	[0.539, 0.674]	[0.468, 0.607]
R2	0.742	0.686	0.521	0.477
R2 Adj.	0.742	0.685	0.520	0.476
Num.Obs.	48473	48453	48471	48470

Notes: TSLS regressions of the probability of 0 income over 30 months for compliers, with block fixed effects using T x 1(Y<.000001) as outcome and d instrumented by the re-centred instrument and block fixed effect, with T = D for Y(1) and T=(1-D) for Y(0). Models estimated jointly by number of children.
We use cluter robust standard error adjusted by block to construct point-wise 95% confidence intervals.

Densities of compliers' individual labour incomes We start by the estimation of the potential individual labour incomes for treated and untreated and compliers. We use data over the same post-treatment period and use κ weights to recover the counterfactual densities. Results are presented in Figure VI.9. In the three panels by number of children; we also added the shape of the in-work benefits stacked over the housing benefits. These have been rescaled to fit the density plots and have no vertical units. They only illustrate the variations of social transfers and marginal tax rate.

Consistent with our previous observations, the estimated densities of treated compliers (blue) show important bunching masses at the *sweet spot*. However, these estimates show that the counterfactual distribution (pink) presents no bunching at all, for any group. Instead, the counterfactual densities are mostly uniform up to the full-time minimum wage, from where they rapidly drop.

The fourth panel shows the aggregate estimates. These results can be compared with the previous bunching estimations with, this time, a more robust causal interpretation. First, there is no difference in density before 50% minimum wage. Then, we have roughly twice as much observations in the sweet spot than in the counterfactual and a missing density from 75% MW until the end.

These estimates show that the programme had strong effects at the intensive margin The most important effect is driven by single parents of two children at baseline. For them, the counterfactual distribution shows no sign of optimisation and is essentially linearly decreasing over the support of the distribution. The estimates for single parents of one child suggest that the programme moved those who would have worked full-time to work part-time, and lowered participation among those who would have worked fewer hours. As for parents of three or more children, we observe higher densities of labour income until the exit point of in-work benefits then a rather linear slope following incentives of the housing benefits.

The illustration of the shape of the in-work benefits over that of housing benefits may help interpreting why working part-time seems to attract participants so much: it also corresponds to the point before they start losing housing benefits. The latter play an important role in these households' budget and loss-aversion may drive part of these behaviours (Fehr and Goette 2007).

Implied elasticities Like in the previous model, we can use the difference in density mass between treated and untreated compliers to estimate the elasticity of labour income that the treatment induced.

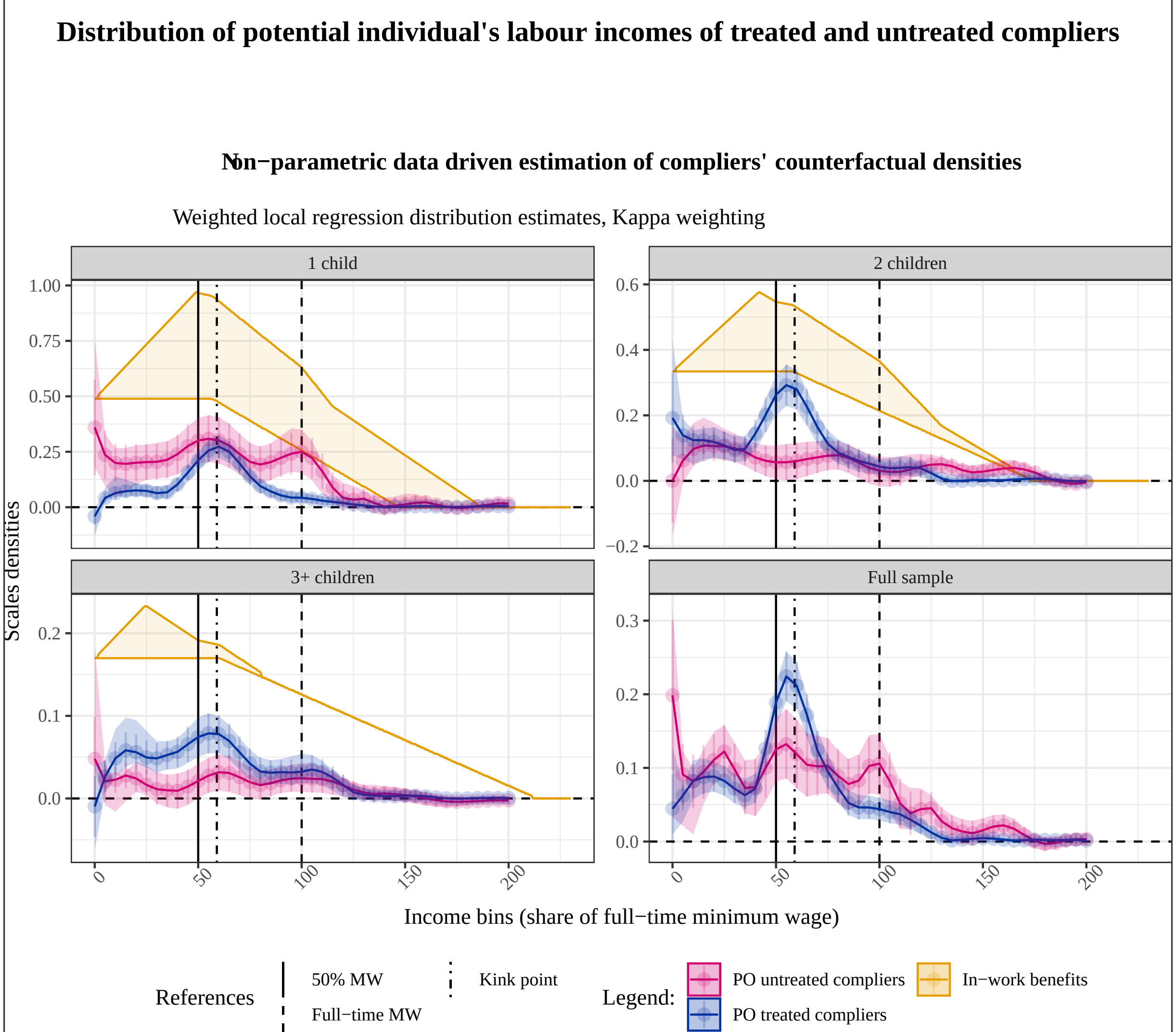
We define the bunching zone as the bins where the joint upper confidence interval of $f_{Y^0|D^1}$ exclude the point estimate of $f_{Y^1|D^1}$ and integrates the distribution over this part of the distribution to recover $\widehat{B} = \int_h^{h+\Delta} \frac{f_{Y^1|D^1}(h) - f_{Y^0|D^1}(h)}{f_{Y^0|D^1}(h)} dh \approx 2.63$, with Y^* the 60% minimum wage kink point. With the average change in tax rate of $dt = 23$ pp, the elasticity can then be approximated as:

$$\hat{e} \approx \log\left(1 + \frac{\widehat{B}}{Y^*}\right) / \log\left(1 - \left(\frac{dt}{1-\tau_a}\right)\right)$$

Which gives $\hat{e} \approx 2.05$, an elasticity 2.3 times higher than the one found earlier with parametric regressions. This new estimate still relies on strong functional form assumptions. However, they use experimental variations and compare the distribution of potential outcomes in the absence of the programme. The difference in bunching mass has a clear causal interpretation.

It is important to understand that the underlying model assumes that the distribution of income is shaped by individual ability a_i, knowledge θ_i and monetised psychological costs ψ_i. In the frictionless model, if the elasticity ε is zero, then the distribution of earnings coincides with the distribution of ability. As ε rises, or if θ_i is shifted to 1 representing *perfect* knowledge, individuals become more sensitive to the tax rate and react by lowering earnings (Bertanha et al. 2023). As a result, the distribution of log earnings becomes a left-shifted version of the distribution of ability. The larger shifts are a consequence of both high tax rate variation and higher awareness of such changes.

Figure VI.9: Counterfactual densities of compliers' own labour incomes



Sources: ALLSTAT, observations from 18 to 30 months since random assignment.
The dependent variable is the individual's labour income as share of full−time minimum wage.
The shaded area above the estimates is the scaled amount of in−work benefits with housing benefits adjusting.
These data are simulated using EDIFIS and rescaled for each panel using the width of the density support.
Notes: Estimations of the potential density functions for compliers following Cattaneo et al. (2021), using the R package 'lpdensity'.
Bandwidth chosen using regularised integrated MSE−optimal and the Epanechnikov kernel function.
Instrument propensity scores estimated by probit of encouragement on block fixed effects to smooth support.
Densities use local polynomial regressions of order 3 , simultaneous 95% confidence intervals use polynomial of order 4 ,
1 degree higher for biais correction.
Densities are scaled over the support of potential income with positive values. The counterfactual probability of
reporting 0 income is estimated by TSLS with block fixed effects using T x 1(Y<.000001) as outcome and d instrumented by
the re−centred instrument and block fixed effect, with T = D for Y(1) and T=(1−D) for Y(0).

Weights are computed for each sub−group separately: $\kappa_0 = \frac{1}{1-D_i}\frac{1-Z_i-(1-\hat{q}_b)}{\hat{q}_b(1-\hat{q}_b)}$ $\kappa_1 = \frac{1}{D_i}\frac{Z_i-\hat{q}_b}{\hat{q}_b(1-\hat{q}_b)}$

The rapid learning and bunching To better understand what changed for treated compliers, we look at the evolution of this density over time, estimating the same model over periods of 6 months. This gives us more plausible proofs that the programme created the bunching mass through reduced optimisation friction and learning of the tax-benefit system We start with the 6 months before random assignment and from 0 to 30 months after and plot the results in Figure VI.10.

This figure shows that the optimisation really stems from the last part of the programme. Before random assignment, distributions overlap and reach 0 before the full-time minimum wage. The recruiting period lasted up to 6 months and during that time, compliers *miss* full-time jobs and we see the lock-in effect starting to show-up on wages of untreated compliers. Heim (2024) showed that the lock-in effect maxed at about 9 months from random assignment. In the window from 6 to 12 months, the counterfactual density gets bigger and flatter while the treated compliers are still barely reporting any income.

In the last 6 months of the programme, almost every new added mass among those with positive labour income bunch at kink point. This bunching mass remains for the next three panels, while the counterfactual density is almost uniform up to the full-time minimum wage. This suggests that some untreated compliers may start part-time jobs then move to full-time, a pattern we do not observe among treated compliers; at least up to this point.

Comparison of untreated compliers with never-takers Since compliers and never-takers may differ in terms of potential outcomes, it is interesting to compare the counterfactual density of untreated compliers $f_{Y^0|D^1}$ with that of never-takers $f_{Y^0|D^0}$. The latter can be directly estimated from observations in the encouragement groups. We estimate these models over the 6 periods to see how these distributions evolve over time and present these results in Figure E.30 in the Appendix.

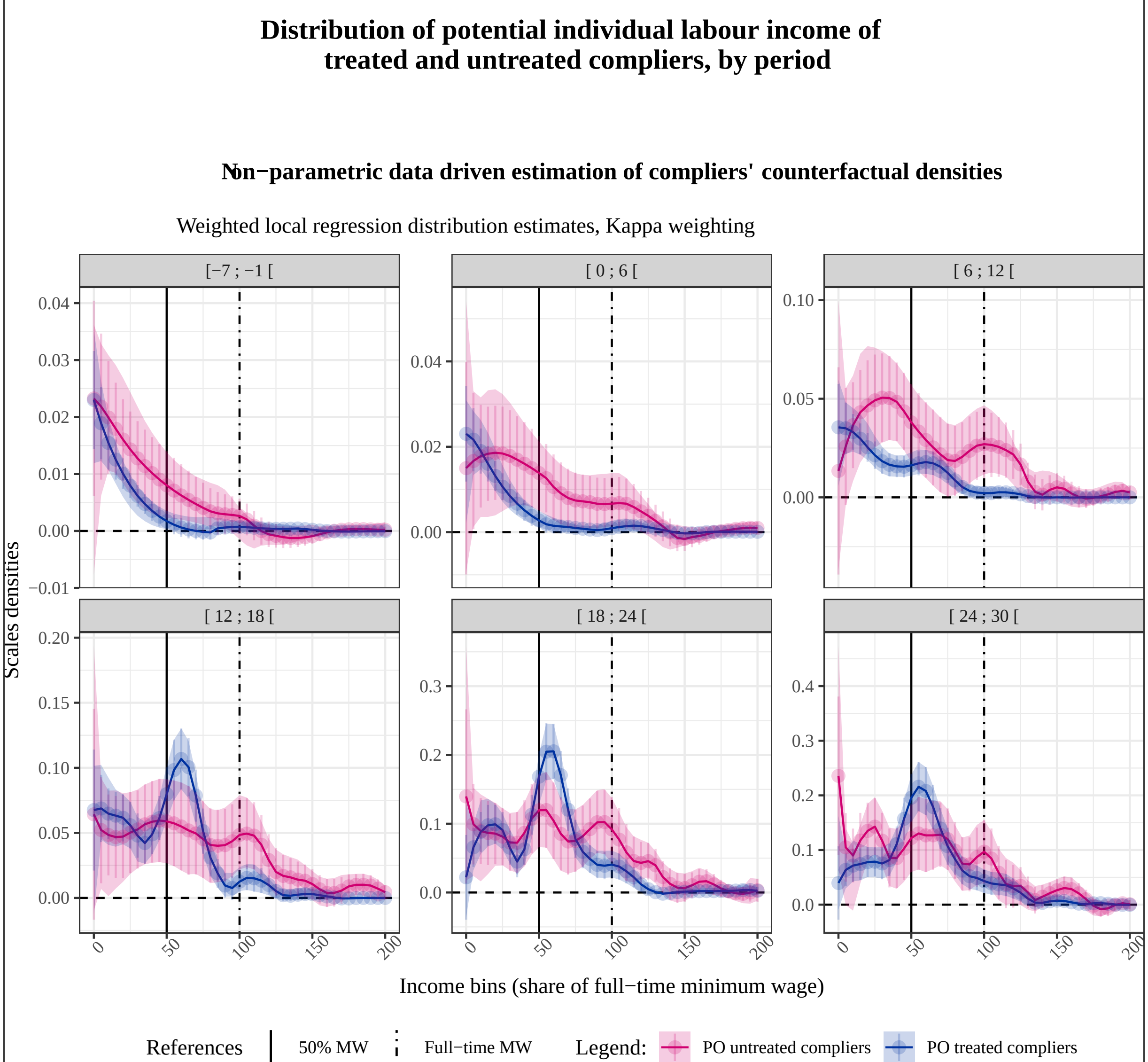
The results show that the potential distributions are very similar in the first two periods, then the lower job finding rates of never-takers leave the distribution lower than that of untreated compliers. However, they do not exhibit much bunching either. While these averages over number of children may still hide differences, they support the idea that in the absence of the programme, households do not optimise much.

Densities of household's labour income for treated and untreated compliers We run similar estimates for household incomes and report them in the Appendix. Figure E.28 uses household's labour income over the same period. Since these incomes include those of spouses, we do not present the shape of in-work benefits for clarity. However, this is the main driver of the amount of in-work benefits.

Unsurprisingly, these estimates are very close to the individual labour incomes, as most remain single parents. The most important difference concerns parents of three or more children. Indeed, the two distributions of treated and untreated compliers largely overlap, while individual incomes were higher among treated compliers in Figure VI.9. For households with three or more children, the programme shifted the source of labour incomes from fathers to mothers. The reported incomes are however significantly lower on the higher part of the distribution. It is also worth noting the shape of the counterfactual density around the kink for parents of two, which exhibits a hole. This further confirms the observation made earlier that untreated compliers would have mostly chosen full-time jobs and not part-time jobs.

We reproduce this analysis over 6 months period to show the evolution of potential household's labour like individual income. The results are presented in Figure E.29 in the Appendix. Results are very similar with one main difference: in the last period, we see the bunching mass dissipate and slightly thicker densities up-to 1.2 full-time minimum income suggesting spouse incomes.

Figure VI.10: Evolution of potential individual's labour income by 6 month periods



Sources: ALLSTAT, observations from 18 to 30 months since random assignment.
The dependent variable is the monthly individual's labour income as share of full−time minimum wage.
Notes: Estimations of the potential density functions for compliers following Cattaneo et al. (2021), using the R package 'lpdensity'.
Bandwidth chosen using regularised integrated MSE−optimal and the Epanechnikov kernel function.
Instrument propensity scores estimated by probit of encouragement on block fixed effects to smooth support.
Densities use local polynomial regressions of order 3 , simultaneous 95% confidence intervals use polynomial of order 4 ,
1 degree higher for biais correction.
Densities are scaled over the support of potential income with positive values. The counterfactual probability of reporting 0 income is estimated by TSLS with block fixed effects using T x 1(Y<.000001) as outcome and d instrumented by the re−centred instrument and block fixed effect, with T = D for Y(1) and T=(1−D) for Y(0).

Weights are computed for each sub−group separately: $\kappa_0 = \frac{1}{1-D_i}\,\frac{1-Z_i-(1-\hat{q}_b)}{\hat{q}_b(1-\hat{q}_b)}$ $\quad \kappa_1 = \frac{1}{D_i}\,\frac{Z_i-\hat{q}_b}{\hat{q}_b(1-\hat{q}_b)}$

Additional estimates In Figure E.26 in the Appendix, we present the estimates on individual pre-tax income using the same method. This outcome includes labour incomes and other non-labour incomes parents report; typically child support or unemployment insurance. For single parents with one child, the programme shifted the entire distribution to the left and participants are also less likely to report small non-labour incomes contrary to the counterfactual. The opposite is true for single parents of three children or more who are more likely to report low levels of income compared with the counterfactual. The estimations for single parents of two essentially overlaps except for a small difference at the sweet spot.

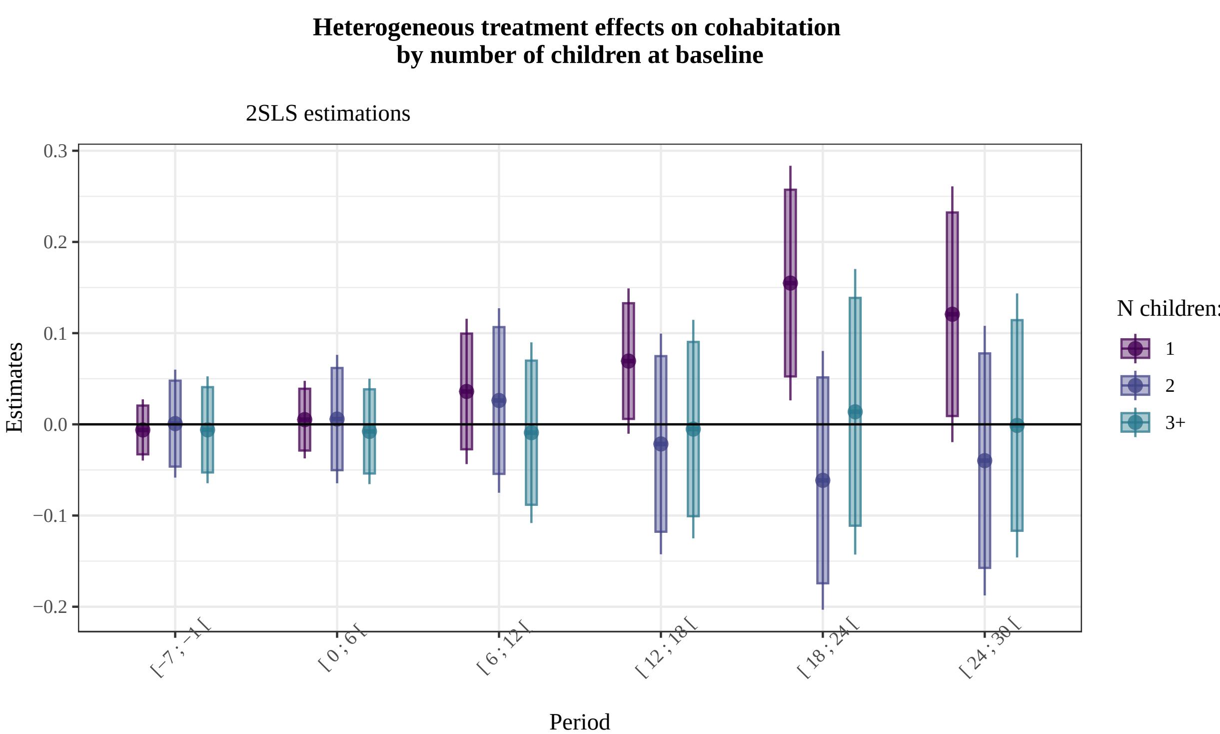
This suggest that untreated-compliers have different source of income, often non labour ones - that treated compliers do not have. The programme dramatically changed the distribution of incomes and shifted them to part-time jobs around the end of the programme. We now look at the effects of the programme on family structure.

3 The effects of the programme on family structure

In this subsection, we estimate the dynamic average treatment effects on the treated by number of children at baseline on the two main outcomes for family size: cohabitation and number of children.

Fewer single mothers with one child at baseline Figure VI.11 reports the estimate of the instrumental variable estimations of the effect of the programme on the probability of cohabitation with a partner by number of children at baseline following the methodology described in V. These results confirm our hypotheses: the programme dramatically increased the probability of cohabitation for single mothers with one child. In order of magnitude, the TSLS coefficient over the year after the end of the programme represents twice the average of the control group. The other groups of parents are not affected by the programme. Note that the fading effects in the end is not due to more break-ups but a catching-up in the control group, as we noticed in the results of Figure IV.7.

Figure VI.11: Average effect of the programme on cohabitation by number of children at baseline

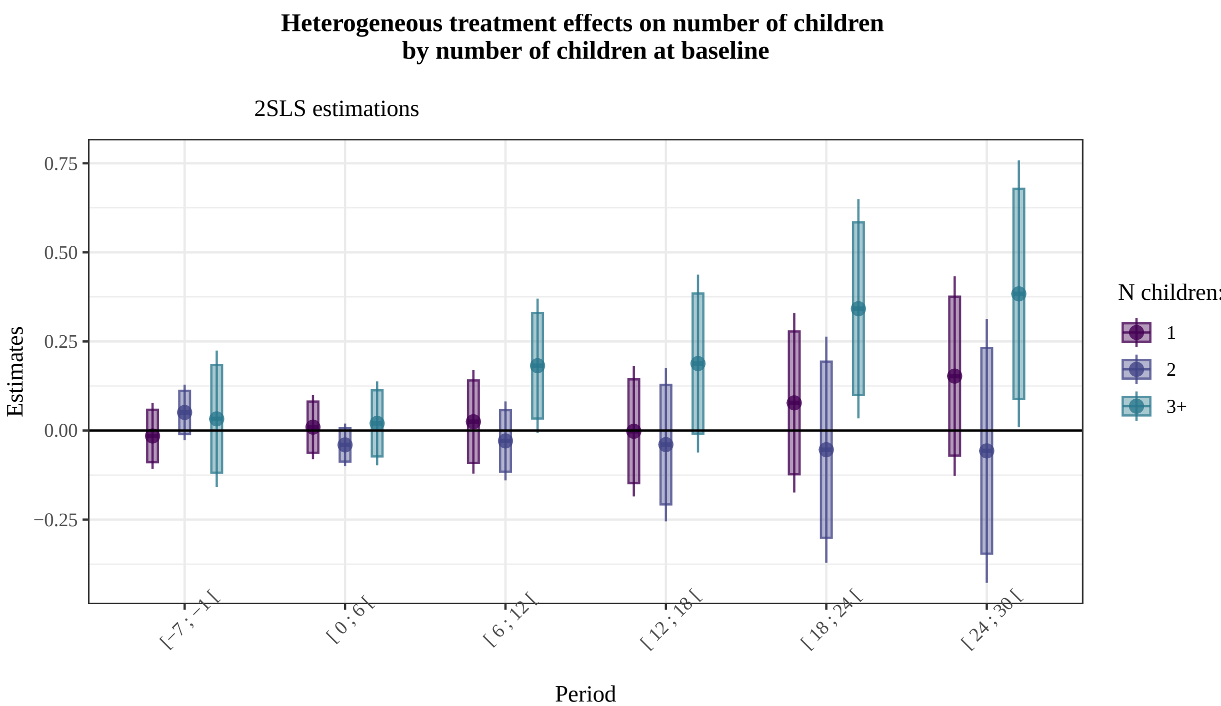


Sources: ALLSTAT 2017−01−01 to 2023−06−01 cohorts 2018 to 2021.
Notes: The dependent variable is the number of dependent children.
Cluster−robust standard errors at the block x cohort level.
− Error bars indicate 95 % pointwise confidence intervals and extended lines account for FWER by subgroup.
− All models include blocks x cohort x relative months fixed effects and instrument propensity score weighting.

These results have to be interpreted with a few caveats in mind. First, the Family allowance fund uses a definition of cohabitation no other public institution formally use. It is different from what the civil code defines and what the National statistic agency uses for households. As discussed in Section 1, recipients must report any change as soon as it occurs, in spite of a very broad definition. We cannot rule out the possibility that the programme increased compliers' reporting of cohabitation - either because of strategic incentives or reinforced fears of punishment - and not *de facto* cohabitation. However, these results are consistent with the incentives we discussed and either show better optimisation or more institutionalised romantic relationships.

Changes in the number of dependent children Similarly, we estimate the effect of the programme on the number of dependent children using the same method and present the results in Figure VI.12. In this plot, we use the definition of dependent children that reflect the RSA and PA updates and present the same estimates with the two alternative measurements in Figure F.32.

Figure VI.12: Relative increase in the number of children for parents of three or more



Sources: ALLSTAT 2017−01−01 to 2023−06−01 cohorts 2018 to 2021.
Notes: The dependent variable is the long difference between the number of children under responsibility at month m and the number of children the month before random assignment.
Cluster−robust standard errors at the block x cohort level.
− Error bars indicate 95 % pointwise confidence intervals and extended lines account for FWER by subgroup.
− All models include blocks x cohort x relative months fixed effects and inverse instrument−propensity score weighting.

These estimates show no effect of the programme on the number of children for parents of one and two children but a very large increase among those who had 3 or more children at baseline. The magnitude of these estimates implies that among this group, almost half of treated compliers have one more dependent children than untreated compliers. While the results with the measure using the definition for family benefits is very close to this one, there is a striking difference with measurement difference based on the definition used in housing benefits (Figure F.32). The number of children in the house sharply increases for those with three children or more from the middle of the programme and continues to increase after.

However, these results do not come from new births but from fewer older children leaving the household. Figure F.31 in the Appendix estimates the same models but instead of splitting the samples by number of children, we estimate separate models by quartile of the age of the oldest children at baseline. The 4rth quartile roughly corresponds to children aged 16 or more at the time of random assignment. The coefficients for this quartile roughly equals the estimates for parents of 3 children, showing that this is the underlying mechanism. For the others, there is no sign of any change in the number of dependent children.

Beyond the presentation of these results and the discussion of their effects on social transfers, it is hard to further interpret what they imply. For sure, they are part of the constrained optimisation problem and may be a way to limit benefit loss. At the same time, de-cohabitation of older children requires some level of autonomy and delayed departure could also reflect changes in higher education or other important life decisions. These data are simply unfit to make further inference.

VII Aggregate effects and additional mechanisms

We run a series of additional estimations to further understand the optimisation we observe and its consequences. First, we analyse the treatment effects of the programme on cumulative labour incomes. To do that, we look at the quantile treatment effect on the annual labour income. Then, we leverage additional variations from the timing of job re-entry to measure the effect of first job re-entry for treated and untreated compliers.

1 Quantile intention-to-treat effects

Our estimation strategy uses **yearly** individual and household labour incomes as dependent variables. As a first step, we estimate the individual sum of incomes over the 12 months after the end of training[51] and define Y^a_{ib} where a is either *individual labour incomes* or *household labour incomes*.

Let $F_{Y^a|Z=1}$ and $F_{Y^a|Z=0}$ denote the distributions of Y^a conditional on being in the encouragement or group, respectively. Then the quantile intention-to-treat effect (QITT)3 is defined as

$$\text{QITT}(\tau) = F^{-1}_{Y^a|Z=1}(\tau) - F^{-1}_{Y^a|Z=0}(\tau)$$

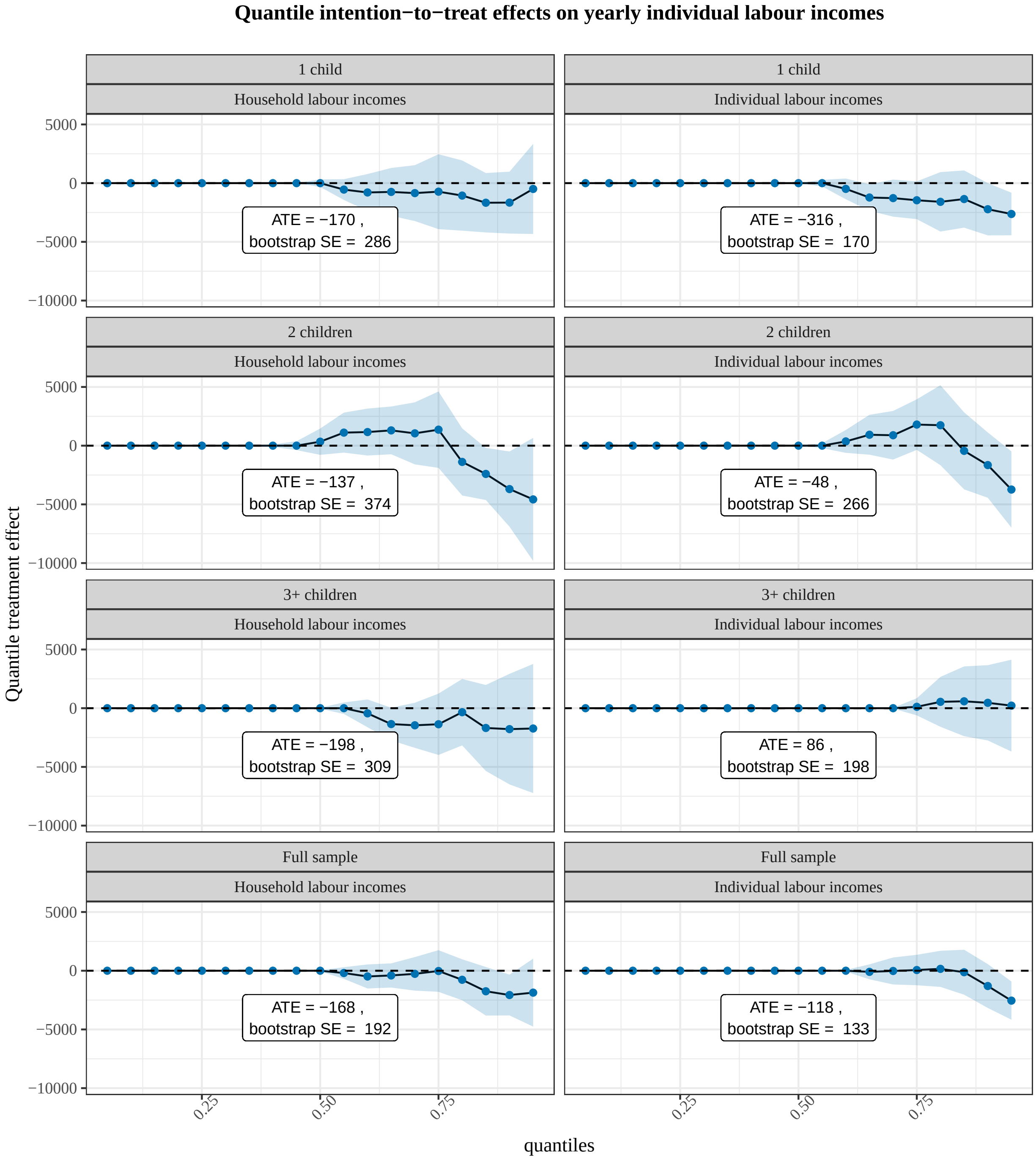
We use the estimator proposed by Firpo (2007) and implemented in the R package `rqte`. This method uses inverse propensity weighting of the distributions and bootstrap standard errors.

Figure VII.13 presents the estimates on the yearly labour income, separately by number of children and for individual an household incomes. Both sets of estimates confirm the previous results, namely that households and individual labour incomes were reduced for single mothers at the top of the income distribution by choosing part-time minimum wage jobs. However, these results bring new insights on the heterogeneity of the responses. Single parents with one child at the top of the income distribution have much lower total individual labour incomes, but the result is less pronounced for their households' incomes, their partners' wages compensating their lower individual incomes. Single parents of two children have pretty much the same quantile treatment effects on their households and individual labour incomes, showing that they are more likely to be the only income in the household. Single parents of three or more children have no significant quantile treatment effects.

We also estimate the same model on individual and household income, including non-labour incomes and report the results in Figure F.34 in the Appendix. They confirm the negative effects at the top of the distribution for parents of 2 and three children, and individual income of those with one child at baseline. However, there is no difference in household income for that group, showing the mitigating effect of increased incomes through cohabitation.

[51] We use $m \in [18:30]$

Figure VII.13: Quantile treatment effects on cumulative individual labour incomes over the year after the end of the programme



Sources: ALLSTAT, observations from 18 to 30 months since random assignment.
Notes: Estimations of the quantile intention−to−treat effect controling for blocks x cohort by inverse−propensity score weightingfollowing Firpo (2007).95% Confidence intervals estimated using bootstrap.

2 The differential effect of job re-entry for treated and untreated compliers

Throughout the paper, we presented participants as better optimiser, strongly reacting to the incentives they learned. The underlying model assumes that they optimise disposable income after tax. The next question is: how does this optimisation affect disposable income per capita ? For that, we then again use the features of our experimental setting to uncover more important causal parameters under minimal additional assumptions. More precisely, we want to estimate how *different* is the effect of job re-entry on disposable income per capita for compliers, had they not participated.

Intuitions So far, we used the variations from individual shocks $\tilde{Z}$ to measure the effect of the programme D on outcomes under 3.1 and 3.2. These two minimal hypotheses ensured by the experimental design could retrieve the full distribution of potential outcomes for compliers aggregating over cohorts. In particular, we can identify $\mathbb{E}[W_{im}(1)|D(1),m]$ and $\mathbb{E}[W_{im}(0)|D(1),m]$, where $W = \mathbb{1}(\text{labour incomes} > 0)$, and the difference between the two is the causal effect of the programme on employment (Heim 2024).

We can introduce a new variable G to denote the month at which household i experience the event "first job re-entry" at month m:

$$G_i = \mathbb{1}(W_{im} = 1, \sum_{m<g} W_{im} = 0) \cdot m$$

Then, $G = g$ indicates that an individual gets their first job re-entry at month $m = g$. Let $F_G(g) = Pr(G_i < g)$ be the cumulative distribution function (CDF) of the timing of first job re-entry. Similarly, let $e = m - g$ denote the timing-of-event.

The main issue is that G is endogenous and censored as many individualq never switch and early switchers may be different from late switchers. For instance, first job re-entry may select different single parents at the early dates than at the end because of unobservable individual endowment, efforts, assortative matching on the labour market and so on. For instance, Figure C.20 in the Appendix presents the distributions of groups G, in number and proportion. We see that the distribution of first job re-entry is downward sloping with a kink around 15 months, $i.e$ roughly the end of the programme. The training period is associated with an increase of first job re-entry among participants which fade out afterward. This creates variations in the *share* of treated compliers among switchers between groups over time.

However, all these problems may actually be solved using the shifts in the share of compliers across groups G, and using the recent results of (Borusyak, Jaravel, and Spiess 2022) on re-centred instruments for *shocks* affecting groups G: the *level* variables. For that, we need an additional exclusion restriction that is: $G(d,z) = G(d)$, the month of first job re-entry does not depend on the instrument beyond participation and can be excluded from potential outcomes. This hypothesis allows us to consider the timing-of-event as an effect of the programme that we can instrument. Since G is a time variable, this additional exclusion implies a parallel-trend assumption, which we can informally assess with estimates on leads of first job re-entry.

With this additional hypothesis, we can recover the potential incomes of treated and untreated compliers around first job re-entry by using i) variations over time of ii) variations of the mass of job re-entry instrumented by iii) variations over time of iv) the mass of compliers.

To gain intuition, first note that under SUTVA and one-sided non compliance, $Y_i = Y_i(1) \quad \forall D_i = 1$ and $\mathbb{E}[Y_i(1)|W_i, e_i = m - g_i, D_i = 1]$ is observed and can be estimated. For instance, we can regress Y on event-time dummies and no constant on the sub-sample of participant and a balanced window of observation around the event. This regression integrates over individuals and groups g observed at month $m = g + e$. This is also true when $e = -1$ in which case $W_i = 0$ and the difference between any month e and $e = -1$. Denote $\Delta_e Y_w(1)$ this long difference:

$$\Delta_e Y_w(1) = \mathbb{E}[Y_i(1)|W_i = 1, e = m - g_i, D_i = 1] - \mathbb{E}[Y_i(1)|W_i = 0, e = g_i - 1, D_i = 1]$$

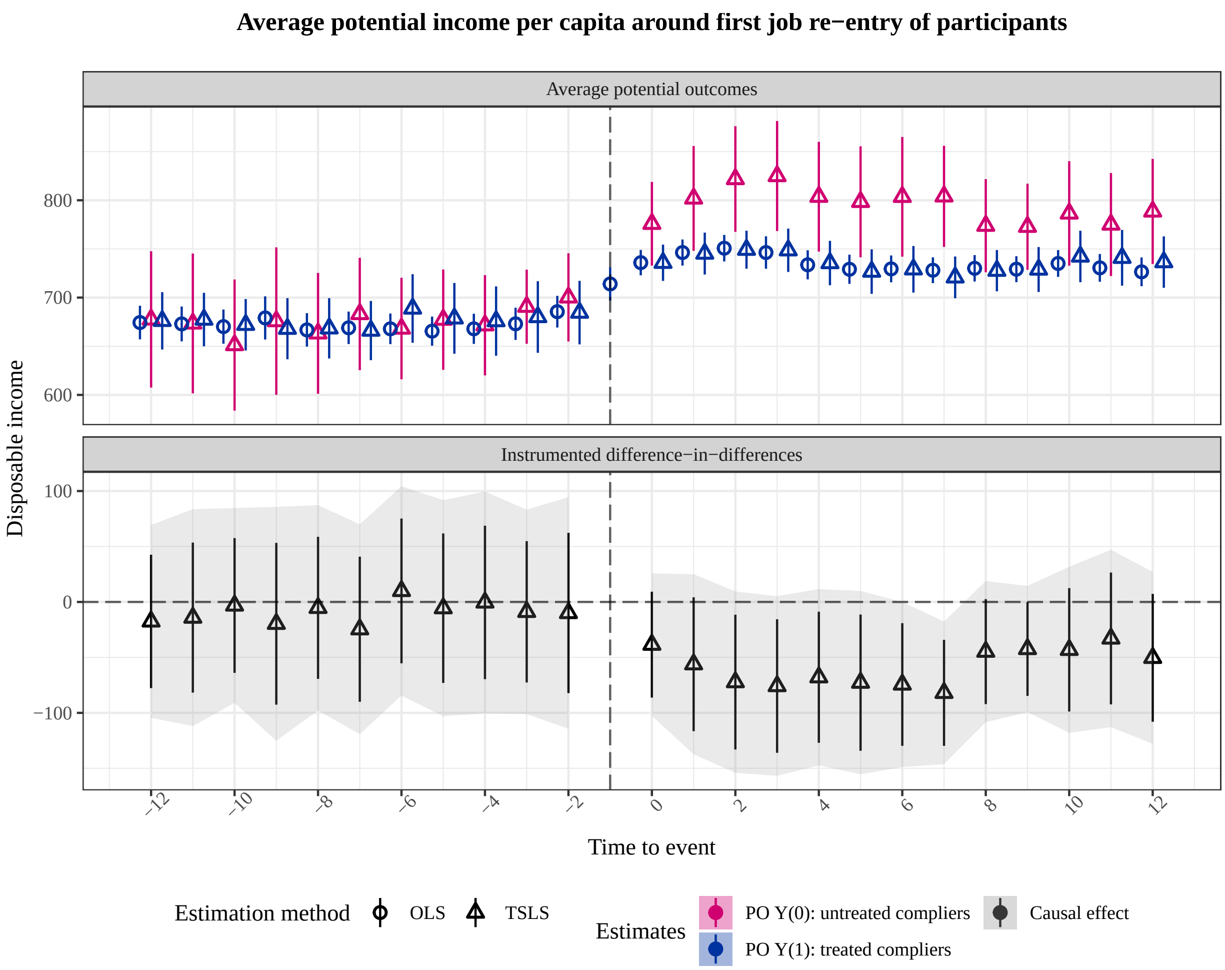
Again, these quantities could be estimated in an event study. However, we can also retrieve the means or their difference using our identification results from Theorem 3.1 and estimate a TSLS event study of $D_i Y_{im}$ on relative-event dummies interacted with D_i and block x month fixed effects, and a first stage of each event-dummy-treatment variables on event-dummy-centred instruments. By replacing D_i by $(1 - D_i)$ in the same equation, we can also retrieve the average potential outcome or average long difference for untreated compliers. Formally:

$$\begin{cases} g(T \cdot Y_{ibm}) = \beta_{bm}\mathbf{B}'S_m(m) + \sum_{e\neq -1} \delta_e (T_i \cdot \mathbb{1}(m=e) + \varepsilon_{ibm} \\ \sum_{e\neq -1} T_i \cdot \mathbb{1}(m=e) = \alpha_{bm}\mathbf{B}'S_m(m) + \sum_{e\neq -1} \tilde{Z}_i \cdot \mathbb{1}(m=e) + \upsilon_{ibm} \end{cases} \tag{3.16}$$

With $T_i = D_i$ to recover potential outcomes of treated compliers, $T_i = (1 - D_i)$ for the untreated compliers.

Each first stage projects the probability of being e months from first re-employment for those with T=1 conditional on the variation in the probability of being e months due to Z conditional on block and month since random assignment. This is where the re-centred instrument makes this TSLS in the framework of Borusyak, Jaravel, and Spiess (2022) and is central for our results to hold. This model uses the variations in the timing of job re-entry caused by the programme - through the lock-in in particular. We present these estimates in the top panel of Figure VII.14. We also display the result of the OLS regression as a validation test. As expected, the two methods yield almost identical estimates. The only difference is precision as the TSLS uses only part of the variation while OLS uses all variations on a restricted sample.

Figure VII.14: Lower potential incomes for treated compliers after job re-entry



Sources: ALLSTAT 2017−01−01 to 2023−06−01 cohorts 2018 to 2021 from 0 to 30 months from random assignment.
Notes: The event W is the first month with positive labour incomes.
IV models for potential outcomes use DY (resp. (1−D)Y) on D (resp. (1−D)) interacted with event−time dummies.
The latter are instrumented by event−time dummies interacted with centred encouragement, with block x month fixed effects instrumenting themselves in the second stage.
The OLS model regress Y on event−time dummies without constant among the sub−sample of participants.
Event−time dummies omit the first month of the window and the month before the event.
95% pointwise Confidence intervals based on cluster robust standard errors adjusted at the block level

From there, we are therefore able to identify:

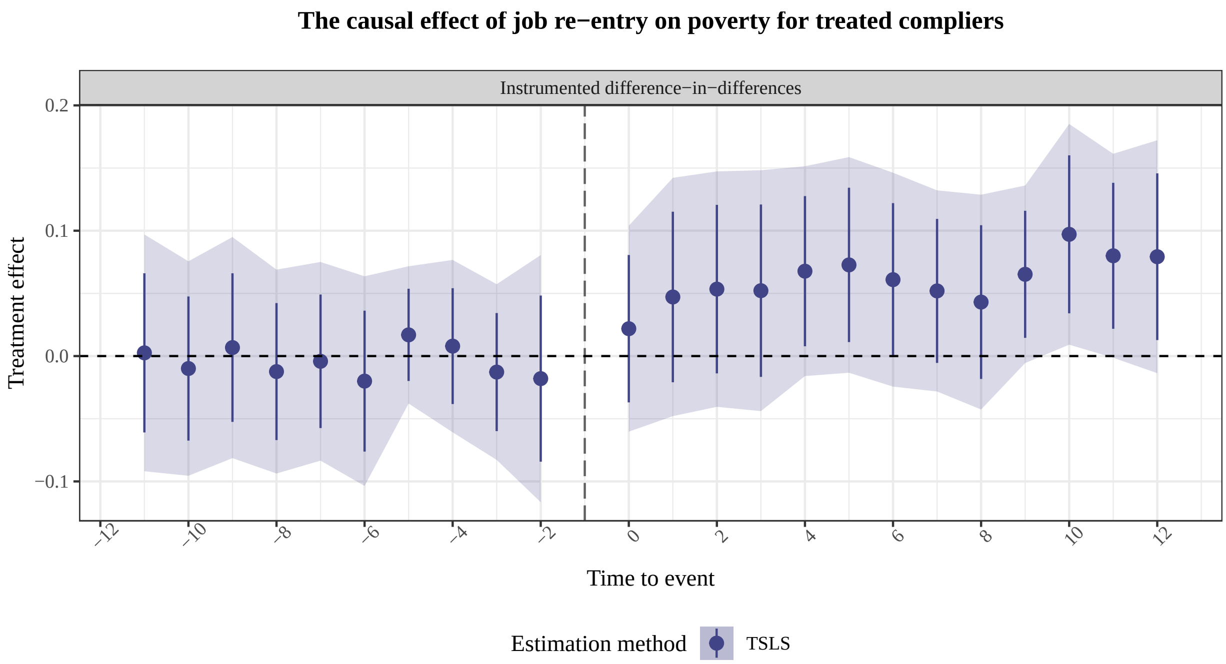
$$\Delta_e Y_w(0) = \mathbb{E}[Y_i(0)|W_i = 1, e = m - g_i, D_i = 1] - \mathbb{E}[Y_i(0)|W_i = 0, e = g_i - 1, D_i = 1]$$

And it is then easy to estimate the triple difference:

$$\Delta\Delta_w(Y) = \Delta_e Y_w(1) - \Delta_e Y_w(0) \tag{3.17}$$

i.e. the difference of disposable income around first job re-entry for treated compliers compared with the effect of job re-entry for untreated compliers. In other words, the differential effect of job re-entry on disposable income due to the programme. These estimates are presented in the bottom panel of VII.14. We build a simultaneous 95% confidence interval using the Holm Bonferroni correction. These results have a causal interpretation under the modified exclusion restriction that the effect of the instrument on the timing of employment only goes through participation.

Figure VII.15: Job re-entry of participants causes higher poverty



Sources: ALLSTAT 2017−01−01 to 2023−06−01 cohorts 2018 to 2021 from 0 to 30 months from random assignment.
Notes: The event W is the first month with positive labour incomes.
The model instrument event−time dummies interacted with participation by event−time dummies interacted with centred encouragement, with block x month fixed effects instrumenting themselves in the second stage.
Event−time dummies omit the first month of the window and the month before the event.
95% pointwise Confidence intervals based on cluster robust standard errors adjusted at the block level.
Simultaneous 95% CI using Holm Bonferroni correction.

Interpretations Figure VII.14 shows that before the first job re-entry, treated and untreated compliers had €676 of income per consumption unit, same for treated and untreated compliers. After the first job re-entry, treated compliers have about € 735 of monthly income per consumption unit, with little variation from the initial shock.

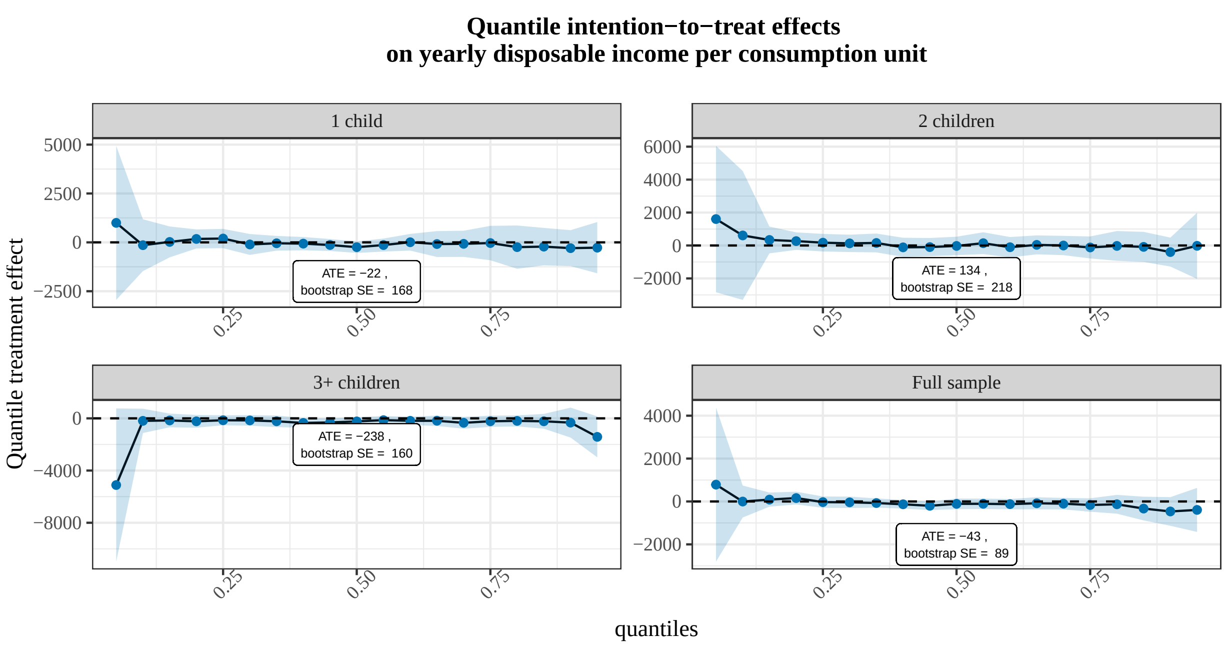
Consistent with our previous results, we see that disposable incomes of untreated compliers are higher after job re-entry than treated compliers, reaching € 796 for the year after. Untreated compliers took jobs with higher wages and treated compliers have lower disposable income when they take a job than they would have had had they not participated. This difference is estimated in the bottom panel and further reinforces the causal interpretation of these results, with absolutely no difference on the lead events. Job re-entry reduces disposable income per capita of treated compliers by € -60 per month on average. Each point estimate excludes 0 from the 95% confidence interval but the joint confidence band are wider and only exclude 0 seven months after job re-entry.

Note that we are not measuring the causal effect of employment as some may switch on and off after. We only measure the difference between treated and untreated compliers stemming from the first job re-entry, and net of other causal paths. Optimisation of labour income at part-time minimum wage reduces the income available for them and their children, but gives them more time out of the labour market than in the counterfactual. However, this lower income also means that for treated compliers, job-re entry increases their probability of living in poverty, as shown in Figure VII.15. This model is estimated exactly like the previous one, only using poverty as outcome.

Overall, the programme reduced participants' labour market participation either through fewer worked hours hours worked (for the most) and at the extensive margin for those with one child. The effect of taking a job on income is lower and increases the risk of living in poverty. However, these are the effects of the programme that go through employment and we showed that there were large effects on family size. Before moving to the conclusion, we want to discuss the fact that while this paper showed large reactions at the intensive and extensive margins, heterogeneous by number of children at baseline, the net effects of all those changes on disposable incomes is precisely 0, on the entire distribution of all groups.

Absolutely no change in the distribution of cumulated disposable incomes In Figure VII.16, we estimate the quantile intention-to-treat effect on the disposable income per consumption unit separately by number of children at baseline. These estimates are strikingly flat: a precise 0 effect on the entire distribution. For these single mothers, optimisation allowed to maintain the same level of disposable income they would have had had they not chosen the programme. The number of children at baseline creates very heterogeneous incentives and except single mothers with one child, they are more likely to be the sole earner. Last, when they manage to receive more money from non-custodial fathers, this amount is 100% deduced from RSA and parents lose family support allowance (See section 3). Recall that at Baseline, only 20.6% of the sample receive child support and 65% perceive ASF instead. Figure F.33 in the Appendix documents the strong effect of the programme on the probability of receiving child support for parents of two and no effect on other parents. However, as soon as the programme ended, the effect dissipates. In the bottom panel, estimates on family support allowance are symmetric. Like those in the control group, they are even less likely to receive the Family support allowance after the programme ended, and as likely to receive child support. Part of this loss of ASF can be due to non-take up or the effect of re-partnering. Indeed, ASF is cancelled when a single parent re-partners, even if the new partner has no link with the children and the father does not pay.

Figure VII.16: Quantile intention-to-treat effects on disposable income per consumption unit



Sources: ALLSTAT, observations from 18 to 30 months since random assignment.
Notes: Estimations of the quantile intention−to−treat effect controling for blocks x cohort by inverse−propensity score weightingfollowing Firpo (2007). 95% Confidence intervals estimated using bootstrap.

VIII Discussion and concluding remarks

This paper analyses the effects of an intensive welfare-to-work programme for single parents implemented in France from 2018 to 2021. Taking advantage of the randomised encouragement design, we study how the programme impacted the distribution and composition of household incomes, as well as family size and structure.

We structured our analysis into three phases. First, we conducted a literature review on welfare reforms and single parents, constructing a theoretical framework grounded in the literature on bunching with imperfect knowledge, psychological barriers, and adjustment costs commonly employed in existing research. However, we critique this approach for its failure to account adequately for the unique circumstances of single parents and question the validity of labour elasticity interpretations that overlook gender norms, children, and household bargaining dynamics. Second, we simulate social transfers by family structure and size using open-source models of the tax-benefit system to estimate implicit marginal tax rates and demonstrate the disparities and divergent incentives. In the third phase, we analysed data from the experiment to quantify the causal effects of the programme on income distribution and family structure. Leveraging the theoretical framework presented earlier, we estimated observed elasticities using parametric models akin to those used by R. Chetty et al. (2011), Henrik J. Kleven and Waseem (2013) or Kostøl and Myhre (2021). However, our main contribution lies in utilising the experimental variations of assignment probabilities to infer the counterfactual distribution of untreated compliers within an instrumental variable framework and measuring effects on family structure. In summary, our main findings are as follows:

1) **Participants bunch at kink points** - between 50% and 60% minimum wage - where the implicit marginal tax rate is minimal, with few reporting incomes exceeding 75% of the full-time minimum wage. This trend is particularly pronounced among single parents with two children at baseline.
2) The distributions of labour incomes among the control group and never-takers exhibit **much lower bunching mass** in the 50-60% range but display another bunching mass at the full-time minimum wage. However, disparities between never-takers and the control group suggest a higher likelihood of part-time work among never-takers.
3) The **implicit marginal tax rate essentially doubles around 60% of the minimum wage**, with variations across households with different numbers of children. At the full-time minimum wage, the implicit marginal tax rate exceeds 70% for single parents with one or two children.
4) The observed elasticities for untreated households range between 0.2 and 0.3, aligning closely with estimates from existing literature. However, the observed elasticity derived from parametric estimations around kink points among participants is closer to 1.
5) Estimations of counterfactual densities indicate that compliers would have reported significantly higher incomes above 75% of the minimum wage had they not participated in the programme.
6) The programme **markedly shifted earning distributions**, reflecting substantial intensive margin reactions, and also affected the extensive margin for single parents with one child at baseline.
7) The elasticity obtained using the counterfactual density of untreated compliers to estimate the bunching mass is approximately 2.05, representing 2.3 times the bunching elasticity found using parametric models among participants.
8) Additional estimations reveal that the effect of the first job re-entry for participants led to a reduction in earned incomes compared to the effect of job re-entry for untreated compliers. Consequently, **treated compliers who secured employment are economically worse-off than untreated compliers who experienced similar job re-entries**.
9) The programme also increased cohabitation among single parents with one child, reduced fertility among single parents with two children, and delayed the departure of older children, primarily affecting parents of three or more children.
10) Ultimately, the programme had no effect on disposable income per consumption unit. The varied heterogeneous reactions and adjustments in labour market participation and family structure resulted in precisely zero effect across all quantiles of the income distribution for all groups.

Internal Validity Our findings are rooted in a well-conducted randomised experiment with administrative data, yielding minimal and balanced attrition. The sample size and rather large first stage with many observations per household contribute to the precision of our estimates. However, it's important to note that our results hinge on the assumption of no direct effect of encouragement on the outcomes of never-takers. While it's plausible that never-takers may have experienced temporary scrutiny and heightened pressure, we posit that any such effect were short-lived and didn't have lasting impacts. Our quantile intention-to-treat analysis, which doesn't rely on this assumption, consistently corroborates our instrumental variable results. We do not use covariates beyond design and time fixed effects and all our models ensure either an average LATE interpretation, or a counterfactual potential outcome distributions or some of their moments.

Employing modern estimation techniques, we derived counterfactual densities from experimental variations, estimated the bunching mass around kink points in the tax-benefit system, and integrated them into a small structural model to compute and interpret elasticities. This methodological approach represents a significant innovation and has revealed notable disparities compared to estimates obtained using conventional bunching estimators. All results point toward the same consistent story: the programme has generated significant changes in participants' choices, extending beyond labour market participation and hours worked to encompass various aspects of household structure and behaviour.

External validity: the effect of the tax-benefit system on poor single parents Acknowledging the inherent limitations of the bunching model employed, it nonetheless provides a framework for understanding the mechanisms through which the programme may have induced substantial behavioural changes. We identified three key parameters potentially affected by the programme: ability, knowledge, and psychological barriers. Did the programme affect abilities ? If we think of abilities as capabilities, then yes. The programme made participants change their behaviours in ways that are consistent with higher agency. However, it is unlikely that the programme affected ability in a way that would shift the distribution of potential outcomes downward without strong elasticities and important frictions the programme alleviated. We argue that lower psychological barriers and better knowledge of the tax-benefit system facilitated by the high-quality support provided are more likely to explain the results we observe. Under these assumptions, the estimated elasticities can be interpreted as closer to structural parameters, suggesting that single parents exhibit high elasticities but face barriers related to knowledge and psychological factors.

Even without these structural assumptions of the bunching estimates, our empirical findings indicate a significant increase in part-time employment among participants compared to what would have occurred had they not participated in the programme. Moreover, our results reveal other substantial behavioural responses that align with the incentives of the tax-benefit system. Cohabitation induces lower loss of social transfers for parents with one child compared to those with more children, resulting in higher rates of cohabitation among this group. Parents of three or more children with low or no labour income rely heavily on housing and family benefits, incentivising them to stay with their children longer, a phenomenon we observe in our data. Despite these adjustments, our analysis demonstrates precisely no effect on disposable income per consumption unit.

If we think that most of the effect of the programme comes from the understanding of the tax-benefit system, our results show that the very high tax burden poor single parent household face creates strong disincentive at the extensive margin for some and a strong economic incentives for part-time jobs in general. Either way, incentives foster situations without enough income to exit poverty and further induce high level of social transfers. In the absence of salient incentives, untreated compliers work more and the problem is not a lack of ability or opportunity, but of tax burden. In the general population, there is already a massive issue of hindered opportunities and 40% single parents working part-time jobs are so involuntarily (Périvier 2022b).

The inadequacy of gender-neutral labour-supply models For a long time, economists neglected the origin of gender differences in the effectiveness of labour market policies. For instance, Card and Hyslop (2005) analyse another randomised welfare-to-work initiative in the USA with 95% women in the sample, which is only mentioned once in the introduction, once commenting table 2 and twice in illustrative examples. The paper develops a theoretical model to explain the low impacts of the programme without considering once the specifics of the sample. Ashworth et al. (2004) produced a meta-analysis of randomised welfare-to-work programmes and also had a sample made of 95 % women, but still generalised their conclusion as if it was balanced between gender. This erases the many constraints women in general, and single mothers in particular, have to deal with while considering their labour market participation. Disregarding childcare, domestic labour and their systemic relegation to women, as well as ignoring the system that pushes them into dependency on either men or the State, can not lead to appropriate or effective programmes - if the goal is to combat poverty through employment.

The literature on taxation also has a very specific language: they explicitly model genderless agents as tax avoiders maximising consumption. We find it ironic that most of the recent literature on bunching was developed using optimisation behaviours of single mothers, acknowledged as heavily stigmatised, while modelling them using such morally loaded terms (Saez 2010; Raj Chetty, Friedman, and Saez 2013). Wording matters. They shape narratives, discourses and political decisions. Eliaz and Spiegler (2020) define a narrative as a "*causal model that maps actions into consequences, weaving a selection of other random variables into the story*". Narratives incorporate normative components in the sense that they make strong yet often implicit assumptions on agents' behaviours. Sometimes, a new terminology is better than a stigmatising short-cut. For single mothers, these representations are both stigmatising and inaccurate, for there is an important missing equation: the domestic production function and more specifically, children human capital [Gelber and Mitchell (2012);Blundell et al. (2016);AttanasioEtAl2018]. These parents chose the level of employment that minimises the time spent in the labour force while minimising the loss of disposable income from taxation, possibly spending more time with their children.

For Maniquet and Neumann (2021), this reaction may very well be an improvement. Focusing on income disregards the labour time it takes agents to earn it. However, labour time is also a determinant of well-being - at least if one defines well-being in a consistent way with preference satisfaction - and an input in the human capital production function of children; a role single mothers in this sample play on their own. When increasing income from below to above the poverty line goes together with an increase in the labour time, anti-poverty policies may decrease the well-being of the income-poor, namely if the latter actually prefers to work less to use their time more efficiently. A large literature comparing time use across different household structures shows that single mothers invest as much - if not more - time in their children's education as coupled mothers[52] (Kendig and Bianchi 2008; Craig and Mullan 2011; Bianchi et al. 2014; Lavoie and Saint-Jacques 2020; Prickett and Augustine 2021).

Tax burden on the poor and coercion: the poor laws of the Twenty-first century ? Our investigation into the French tax-benefit system has unveiled significant disparities and disincentives for labour participation, particularly among single parents. In particular, single parents on welfare receive no child support. The latter is 100% taxed through lower welfare payment, reducing by the same amount any monetary incentive for employment unless the parent earns enough to be ineligible for in-work benefit. Then only would they receive child support. All other family benefits are also fully taxed for welfare recipients while they are not part of the taxable income for parents out of welfare. Moreover, single parents, especially those with multiple children, face disproportionately heavy tax burdens, particularly at income thresholds near minimum wage levels. This situation creates significant disincentives for increasing labour participation and hampers efforts to improve disposable incomes. The intricate interactions between various social transfers create a web of complexities that further exacerbates the tax burden on vulnerable groups. This complexity contributes to a lack of transparency and understanding among beneficiaries, hindering their ability to make informed decisions about employment and financial planning.

We coined the word *"Assistaxation"* to convey the idea of providing assistance in a way that is burdensome, overly taxing, or difficult mentally, physically, emotionally, or financially. It is not just about the administrative burden and stigma discussed in the literature; it also involves heavier and implicit taxation of labour income and a 100% tax rate on child support and family benefits of the poorest. *Assistaxation* also better reflects what type of taxation is being avoided here: sophisticated scheme and coercive measures to force labour market participation while taxing

[52] A notable fact from Dotti Sani and Treas (2016) is that French parents are the only ones in their international comparisons who spend less and less time with their children.

them most unless they leave assistance or repartner. If we were to make an optimal taxation argument considering the high elasticities we found, we would typically go in the direction of a gender-based taxation accounting for these structural differences in elasticities, as studied by Alesina, Ichino, and Karabarbounis (2011). For the current system sets monetary incentives such that the optimal level leaves working single parents and their children in poverty, while still relying heavily on social transfers.

The practices of the CAF have faced criticism and condemnation from the "*Défenseur des droits*" (Défenseur des droits 2017) in France, particularly in cases of perceived harassment towards economically vulnerable households. Recipients often report a lack of notification, discovering deductions on their CAF personal accounts or bank statements without understanding the rationale or procedures. These deductions can be substantial, reaching up to 50% or even 100% of the recipients' resources, raising concerns about their legality and fairness. The scrutiny exercised on vulnerable individuals, according to Défenseur des droits (2017) is frequently deemed disproportionate, discriminatory, and often lacks legal foundations. The control agents, except those from the Employment agency, have the authority to request various documents directly from banks, employers, energy providers, telecommunications operators, etc., without being bound by professional or banking secrecy. The scoring system used to trigger these controls, relying on the absence and variability of income, tends to disproportionately target the poorest beneficiaries (Quadrature du net 2023). As indicated in the same report, the probability of an RSA beneficiary undergoing controls was 40 points higher than their proportion in the overall population receiving social benefits.

Drawing parallels with the historical context of the Poor Laws reveals striking similarities in the systemic inequalities and barriers to economic advancement faced by marginalised groups (Persky 1997). Just as the Poor Laws entrenched poverty and perpetuated social stratification, "assistaxation" perpetuates cycles of financial hardship and inequality, hindering efforts to achieve economic security and social mobility.

From "deserving Poor" to "Welfare Queens": a backlash against women In her work titled "Backlash Against Welfare Mothers Past and Present," Reese (2005) analyses the intersection of class, race, and gender issues during conservative political periods. Reese (2005) demonstrates political alliances rooted in racist sentiments and patriarchal family values. Whether it's the TANF (Temporary Assistance for Needy Families) in the USA or allowances for single-parent families in the British New Deal, these reforms reinforced the "ethics of marriage," the idea that poor women should marry and remain married to a man. State governments and local administrations commonly used "suitable household" and "suitable mother" rules to deny public assistance to single mothers (Fisher and Reese 2011, 231). The narrative of *"Ending welfare as we know it"* of President Clinton has been highly pervasive and led to a political evolution of the concept of the "deserving poor"[53] (Peterson 1997; Carcasson 2006; Ellwood 2000). In her doctoral thesis, Mangin (2021) demonstrates that the narratives used in debates accompanying welfare reforms are reflected in the legislative corpus of different States, including stigmatising stereotypes like the "*Welfare Queen*", which depicts single mothers living on welfare as manipulative, dishonest, unworthy, lazy, African American women having children to avoid work and challenging dominant sexual norms and gender roles (Jarrett 1996; Foster 2008; Van Doorn and Bos 2019). These stereotypes are also found in the discourse of some social workers and are reflected in their daily interactions with these beneficiaries (Masters, Lindhorst, and Meyers 2014). Similarly, in Great Britain, the image of the irresponsible white working-class single mother who became pregnant at seventeen and had children with multiple men was continually reinforced by government rhetoric and media representation (Herke 2021; Herbst-Debby 2022).

In France, several recent work have documented the large inequalities and unfair treatment of single parents (Le Pape and Helfter 2023); some even proposed and simulated various reform scenario (Allègre, Périvier, and Pucci 2021). Perhaps the most immediate reform would be to remove child support and family support allowance from the income deduced from welfare and in-work benefits, as proposed by Pucci and Périvier (2022). Our contribution further documents the lack of transparency of the system and its impact on single parents, along with activation measures. Single mothers' unfair situation is unlikely to improve so long as the fiscal, social and administrative reference centres on the idea of couples while households get more and more diverse, often falling in between definitions.

[53] See the works of Appelbaum (2001) and Robert A. Moffitt (2015) regarding reforms in the United States or Herke (2021) for a study of the Hungarian case compared to other European Union countries

Appendix

A Details on the French tax benefit system

A.I A short history of monetary-incentives in France

A minimum income scheme from the 1980's In 1985, France introduced a welfare benefit called *Revenu minimum d'insertion* (RMI). The RMI was a differential minimum income imposing a 100% implicit marginal tax rate. When one got a job, a temporary reduction in marginal tax rate was implemented[54] such that if a recipient took a job, only half of their labour income was considered to compute the level of RMI. The reduction was only implemented for the first 750h worked. Gurgand and Margolis (2008) analysed monetary incentives using data from 1996-1998 on recipients of RMI and consider the distribution of potential monthly earnings based on observed wages and working time. Accounting for the welfare earnings top-up programme (*intéressement*), the study finds that gains are almost always positive but generally low, particularly for single mothers. Using census data from 1999, Bargain and Doorley (2011) analyse the causal effect of RMI on employment level of childless single men using the 25-year-old eligibility threshold and show large negative effects. No such effects are found for single mothers.

A first tax credit in the early 2000's Inspired by optimal taxation theories from the US, in-work benefits started being introduced in the French socio-fiscal package in 2001 with the *Prime Pour l'Emploi* (PPE) or Job Bonus. The stated goal was to create a monetary incentive targeting poor workers, all the while combatting the risk of "inactivity traps" created by the *Revenu minimal d'insertion* (RMI), a basic income scheme. Evaluations of PPE quickly underlined that the policy wasn't meeting its goal. In 2002, Cahuc (2002) ascribes this failure to four factors : (1) employment is constrained by demand because of the minimum wage's high value ; (2) incentives for full-time re-employment are already plentiful, conversely to part-time incentives ; (3) while some jobs have been created, PPE incites some women to reduce their working time, since it is degressive above minimum wage and includes the household's incomes in its calculation ; (4) the fact that those benefiting the most from PPE are households belonging to deciles 2 to 4, and not the poorest households, makes its impact on poverty very weak. Stancanelli (2008), utilising various double-difference strategies and data from the Employment Survey, found a negative effect of the PPE on married women's labour market participation (-3 percentage points), a positive effect on cohabiting women (+6 percentage points), and a negligible effect on single women[55]. Aggregate effects on women's labour market participation were modest (around 2,000 job entries). Despite an average take-up of 95%, Vermare et al. (2008) did not find any significant effect of the PPE on employment in the general population, including married women[56].

The activation turn of 2008 In 2007-2008, the in-work benefits has been tested in an experiment and generalised before results were known in 2008 (Bourguignon 2009). The new welfare replaces the former basic income RMI and the temporary allowance for single parents API (*Allocation parent isolé*), replaced by a similar temporary RSA-supplement (*RSA majoré*). Conversely to the report's recommendations, in 2008 RSA and its in-work benefits (*RSA activité*) were adopted[57] with a flat implicit marginal tax rate of 38%, with no difference between part-time and full-time. Another deviation from the report is that the degressivity rate no longer depends on family configuration, which favours families with children (Allègre 2024). To be eligible, one must be at least 25 years old, reside in France, and not be enrolled in school. Recipients are required to actively search for employment and participate in social or employment support programmes, or risk facing sanctions. Temporary supplements for single parents are accessible for 12 consecutive months or up to 18 months from the triggering event. Single parents with *RSA majoré* must also participate in job-search or social support activities unless their youngest child is younger than 3. For such households, there is no age requirement or mandatory activation and RSA-supplement extends until the youngest child turn 3.

[54] Called *intéressement*, top-up programme in the paper.
[55] The negative impact on married women was primarily attributed to a resource condition based on couple income.
[56] They used data from the Tax Income Survey (ERF). Note that the high take-up rate is due to the easier access as it was embedded in tax forms. See Allègre (2024) for a discussion.
[57] While political opposition prevented RSA activité from being financed by the suppression of PPE, the latter's weight de facto diminished due to the freezing of its scale : in 2014, its average amount was 33 euros monthly (Allègre 2024).

Regarding the effects of the in-work benefits (*RSA activité*) introduced in 2009, the improvement in monetary incentives did not necessarily translate into increased employment, as the initial evaluation showed (Bourguignon 2009). Most research find little to no effect in general, but some effects when considering sub-groups. Simonnet and Danzin (2014), using administrative data from Cnaf and a double-difference strategy, observed a positive impact on the employment of single mothers, but not on men. Allègre (2011) showed a negative and significant effect on the labour supply of married women but a positive effect for single women. Bargain and Vicard (2014) used the same identification strategy as Bargain and Doorley (2011) with a notable difference: they looked at 'the youth' and pool genders together. They conclude that the disincentive results of the RMI were weak and concentrated among low-skilled workers, and find no disincentive effects for RSA. While the main results of Bargain and Doorley (2011) was about heterogeneous gendered differences, this analysis did not even report separate estimates by gender, nor between parents and non-parents. Sicsic (2019) explored the elasticity of earned income concerning marginal tax rates in the French socio-fiscal system, revealing a modest elasticity of 0.1 in response to changes in RSA marginal rates.

Reducing welfare stigma through a separate in-work benefits In spite of several issues, the need for a new reform arose outside of economic concerns : while non-take up of RSA approximated 40%, it reached a staggering 68% for RSA activité (Domingo and Pucci 2014). Misinformation played a role (people fear having to reimburse overpayments) but explanations primarily underline the power of stigma: the in-work benefits was associated to RSA, both by name and administrative process, thus making it *repugnant* for poor workers averse to a "hand-out". A report by MP Christophe Sirugue further notes that generally, "*RSA activité and PPE's impact is very weak when it comes to incitations to either (re-)entering or remaining in employment*" (Sirugue 2013).

The 2015 reform - implemented in 2016 - aimed both at combatting this repugnance and at simplifying the social benefits "menu" by merging together RSA activité and PPE into a single in-work benefits : *Prime d'activité* (PA). Just like *RSA activité*, PA degressively declines (by 39%) with earned income, and its amount is the difference between a base amount and the household's ressources. Unlike PPE however, it also includes a progressive individual bonus for every household member earning more than half the minimum wage. While this reform has been analysed as discretely shifting the goal from incitation to distribution (Allègre 2024), it is also worth noting that its budget was based on an estimated take-up of 50% (or a rough average between the two previous schemes). In reality, it reached about 70% as soon as the first year, leading to annual additional expenses of 800 millions euros.

RSA and PA remain tightly linked in their formula and eligibility rules. The main difference is that if a household applies for PA with incomes lower than the RSA, but does not receive RSA already, they have to fill another form for RSA[58]. Conversely, PA is automatically computed for RSA recipients who report positive labour incomes. PA is also open from 18 years old while RSA is still limited to adults older than 25, except single parents[59].

A massive increase of in-work benefits in 2019 The 2019 so-called "Gilets Jaunes" reform was presented as a concession following the eponymous social movement that had shaken France for months, with a notable focus on poor workers' declining purchasing power and living conditions. We provide a detailed explanation of the PA formula and the effect of the reform in Appendix A.III that follows.

While the reform had an explicitly distributive goal, with President Macron promising a € 100 increase of PA's individual bonus, many interpreted it as more of a publicity stunt. Indeed, out of the 100 euros promised, 10 were expected due to the minimum wage's automatic revaluation, and the rest was already planned as early as 2017, initially through successive annual increases (Bozio et al. 2023). A more significant change resides in the enlargement of income eligibility levels by the inclusion of incomes up to 1.5 minimum wage (vs 1.3 previously). Using microsimulation on administrative data from CNAF, Dardier, Doan, and Lhermet (2022) estimate that 83% of the subsequent 43% take-up increase are due to newly eligible beneficiaries, while the remaining is ascribable to the reform's visibility. They also estimate that the reform reduced poverty by 0.6 point, with most gains located in households of the 2nd and 3rd living conditions decile. For families with children, including single parent families, the poverty rate reduction is over 1 point (Dardier, Doan, and Lhermet 2022). It is note-worthy that a majority

[58] something many don't know which largely contributes to non take-up of RSA, as documented by Hannafi et al. (2022).
[59] The age threshold allowed Locks and Thuilliez (2023) to demonstrate that RSA reduces homelessness and estimate that the cost of setting the age limit to 18 would be 60% offset by savings in social assistance costs to the homeless.

(57%) of PA beneficiaries are women, which aligns with their over-representation in part-time employment, and raises questions as to the efficiency of incitations at the intensive margins when they are blind to gender.

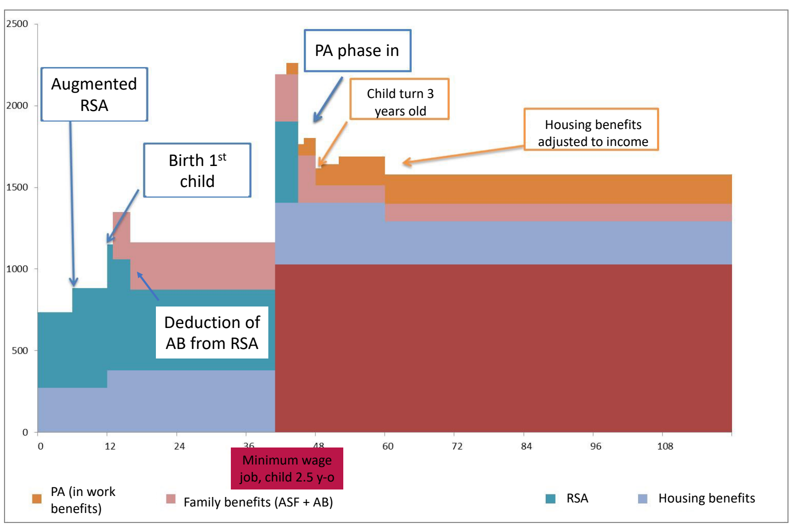
So far, we are not aware of a successful causal evaluation of the effect of this reform. The latter being highly confounded by mass public comunication, it is hard to disintangle the reaction to the reform from better knowledge of the tax-benefit system. Bozio et al. (2023) tried to estimate the effects of the reform using a difference-in-differences design comparing groups most affected by the reform with groups less incentivised. Using micro-simulation and case-studies, they predict different gains for households varying by family size and composition. However, a placebo test on the year before the reform between these groups finds similar estimates and reject the parallel trend assumption.

In summary, French evaluations of the effects of monetary incentives suggest positive effects for singles and single mothers, with relatively limited magnitudes. Effects on other groups are either null or negative on average. While which effects dominate remains essentially an empirical question whose answer is context-dependent, the challenge also stems from identifying and estimating the intensive margin reaction and the effect on wage.

A.II Dynamics of social transfers

An overlooked aspect of the research on the effect of social transfers are their schedules: how and when benefits or tax adjust. While it matters a lot for incentives, the dynamics of adjustment of various social transfers when people's situation change is an under-investigated topic. The main reason may be that nobody really knows how it actually works. Let-us take a simple case-study to illustrate what happens when a single mother gets pregnant and returns to work when her child reaches 2.5 year old.

Figure A.17: Dynamic simulation of the effect of job re-entry at the minimum wage on social transfers in 2018 for a single mother when her child is 2.5 years old. Old internal DSER simulation tool.



Up to 2018, the statistics department of CNAF used to build what was meant to be a simple case-study simulation tool to model the adjustment to job re-entry in particular. Figure A.17 displays a simulation using the 2018 legal framework for the aforementioned example. Pregnancy starts at $m = 0$, which coincides with her last month of RSA in the quarter - an important detail we discuss right after. Three months later, her physician declares the pregnancy, which triggers the *augmented RSA* in the next quarter and until the child turns 3. When the child is born, housing benefits increase and two more transfers are opened: the *Family support allowance* (ASF) and early childhood benefits[60] (AB). At birth, parents may also receive a Birth allowance[61], a one-time transfer to help with higher

[60] AB stands for *allocation de base* which is the amount of cash benefit for low income families with children under 3. Like any other social transfers, they are deduced from the baseline amount of RSA and PA.

[61] Included in the augmented RSA here. It targets both single and coupled parents, with a different income threshold of eligibility

spendings around childbirths. In the next quarter, 80% of ASF and the entirety of family benefits are deduced from the augmented RSA. ASF is perceived because we assume that the father does not pay child support. Otherwise, the amount of child support would be entirely deduced from RSA. In this simulation, the mother takes a full-time job when the child turns 2.5, again on the last month of the quarter, and immediately reports her first wages. In this first quarter of employment, she cumulates the full RSA payment with her earned income, and the PA slowly kicks in since in her next quarterly report, she has only been working for a month. The next quarter, she loses the RSA - a full-time minimum wage being higher than the threshold - and the PA increases. Then the child turns three and there are no more early childhood benefits. The reduction of the early-childhood allowance causes an increase of the in-work benefits. Indeed, since social transfers are part of the reference income, they are entirely deduced from RSA and PA payments and their exhaustion is partially offset. The last adjustment occurs two years later when the taxable income of the first year in employment is available and housing benefits are reduced[62].

This figure shows that there are monetary incentives for employment but their adjustments are highly dependent on many other parameters that together, may very well create situations that are hard to track for parents, especially since no detail on computation is provided. This tool has been discontinued due to the challenges of keeping it updated with frequent reforms, changes, increases in minimum wage, and other factors. Moreover, the tool faced difficulties in presenting typical scenarios due to the intricate consequences of any alterations in the timing of events. In the given example, even a minor shift in the month of job re-entry within a quarter significantly impacts the distribution of in-work benefits over two quarters. The intricacies are further heightened by precarious employment situations[63].

In the next appendix, we present the details of the formula for the in-work benefits and the main changes of the 2019 reform.

A.III The in-work benefits (PA): formula eligibility and the 2019 reform

The amount of in-work benefits PA_m for the month m, is computed according to the following formula[64] :

$$PA_m = \underbrace{BPA_m\left(1+\delta_m^f\right)}_{\text{Flat amount}:FPA_m} + \tau_m \tilde{W}_m - \underbrace{\max\left(BPA_m\left(1+\delta_m^f\right), \tilde{Y}_m\right)}_{\text{Reference household incomes}} + \underbrace{\sum_i B\left(\bar{W}_{iq}^a\right)}_{\text{PA supplement}} \tag{3.18}$$

In words, the amount of in-work benefits is based on the difference between a flat amount and households' incomes to which an individual bonus is added. The flat amount of the *Prime d'activité*, denoted as FPA_m, depends on the current baseline amount BPA_m multiplied by a factor δ_m^f, weighting family composition of the household. BPA_m represents the amount for a single individual without children, and δ_m^f is a weighted sum of household members where the one additional person (child or partner) means $\delta_m^1 = .5$. An additional second person leads to an additional 30% increase in the amount, 40% for single parents. Any person beyond three additional individuals qualifies for a 40% increase. To this flat amount, a fraction τ of the household's earned income $\tilde{W}_{im}$ during month m is added, and the household resources $\tilde{Y}_{im}$ are subtracted[65]. Note that the household's earned income $\tilde{W}_{im}$ is included in the total resources $\tilde{Y}_{im}$ with spouse's incomes and other resources. Importantly, other resources include all other social transfers, such as child support or capital incomes (if any). Moreover, it is the sum of earnings that matters and the composition of incomes between spouses has no influence on the amounts but through the individual supplement. The French administration institutionalised an old-school Beckerian static unitary model (Périvier 2012), as if the household's preferences can be represented using a unique utility function that does not depend on prices, incomes, or any exogenous factor, independently of the number of household members (Pierre-André Chiappori 2017).

The bonuses are calculated based on the average earned incomes $\bar{Y}_{iq}$ of each member i of the household during the reference quarter q. Parameters BPA_m, τ_m, δ_m^f are set by decree and have evolved over time, using the following formula[66]:

[62] Since 2021, this is no longer the case. The current system defines the reference income from a moving average over earned incomes in the past 12 months.
[63] For instance, households who work for a single quarter will have very different in-work benefits depending on which month they first registered.
[64] *Article L842-3 du Code de la sécurité sociale*
[65] The resource base is considered to be at least equal to the flat amount, hence the max function.
[66] *Articles L842-3 et D843-2 du Code de la sécurité sociale.*

$$B\left(\bar{W}_{iq}\right) = min\left(\bar{B}_m, max\left(0, \bar{B}_m \times \frac{\bar{W}_{iq} - S_{min}}{S_{max} - S_{min}}\right)\right) \tag{3.19}$$

To understand this formula, let me describe which parameters are set by law and how they define these variables. The legislation defines three legislative parameters by decree regarding the calculation of individual bonuses:

- The first parameter, τ_b, corresponds to the maximum percentage amount of the bonus relative to the base amount. The maximum bonus amount, expressed in euros, is then calculated as $B_m = \tau_b BPA_m$.
- The second parameter, s_{min}, corresponds to the minimal threshold of monthly earnings (in multiples of the gross hourly minimum wage) required to qualify for the bonus. The earnings threshold in euros is thus $S_{min} = s_{min} \times \bar{s}$, with $\bar{s}$ being the amount of the current gross hourly minimum wage.
- The third parameter, s_{max}, corresponds to the threshold of monthly earnings (in multiples of the gross hourly minimum wage) from which the bonus becomes maximum and constant. The earnings threshold in euros is therefore $S_{max} = s_{max} \times \bar{s}$, again with $\bar{s}$ being the amount of the current gross hourly minimum wage.

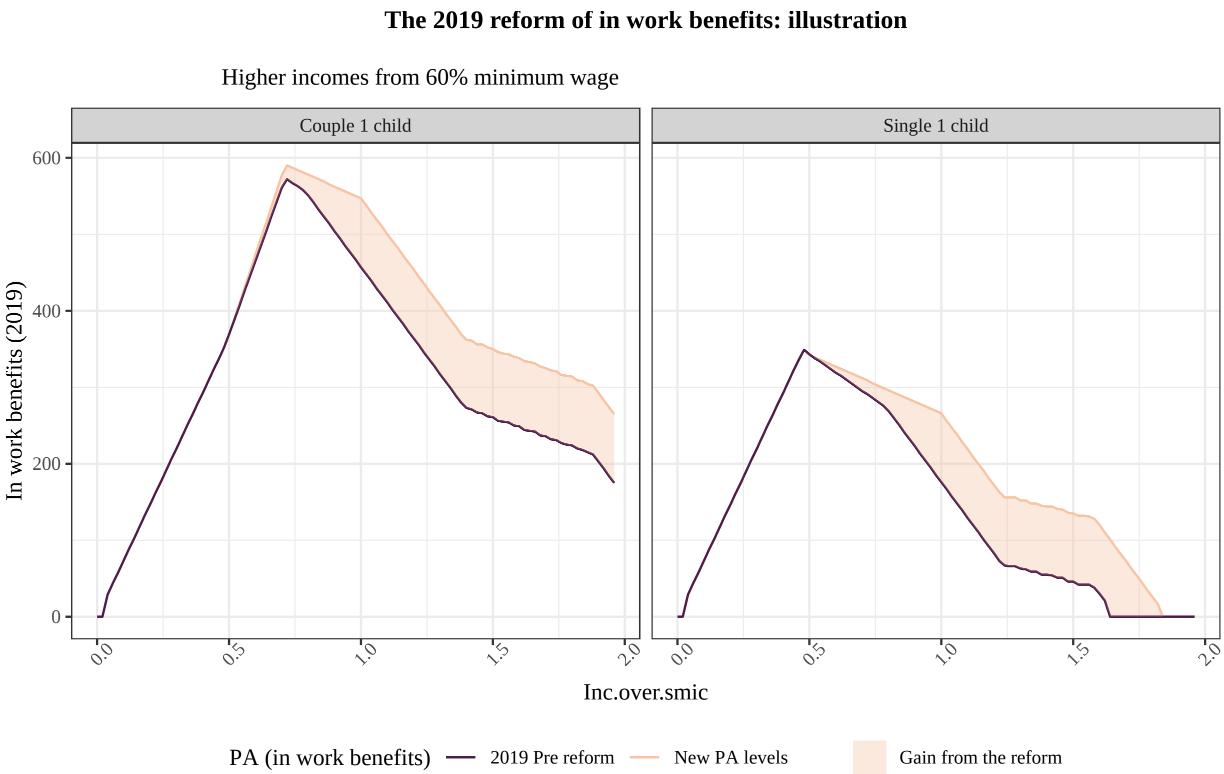
The bonus amount for which an household is eligible depends on each individual's average income over the reference quarter. It is zero when these earnings are below S_{min} euros. It then increases linearly with the amount of these earnings until it reaches the maximum amount, B_m, when the individual earned incomes equal $\bar{B}_m$.

In January 2019, following the "Yellow Vests" movement, a major reform affected three components of this formula:

- Higher baseline amount BPA_m,
- Higher PA bonuses
- Changes in the marginal tax rate on earned incomes, increasing or decreasing depending on the configuration.

An important effect of the reform is the alignment of the reduction in individual bonus at the minimum wage level. Figure A.18 reproduces data from Dardier, Doan, and Lhermet (2022) who used microsimulation to measure the effect of the reform. They estimated that this reform would have led to a 37% increase in the number of households benefiting from the in-work benefits and an average gain of €70 per month at an increased total cost of nearly €4 billion.

Figure A.18: Effect of the 2019 reform of in-work benefits on transfers for parents of one child as simulated by Dardier et Al. 2022



Sources: Dardier Et Al (2022), microsimulation
of the reform for household with one child, couple with one earner.

A.IV Variations in the implicit marginal tax rates

Ultimately, the 2019 reform induced large changes in the shape of the in-work benefits and introduces more variations in the marginal tax rate. To see that, we use the EDIFIS simulation model of DREES and compute the implicit marginal tax rate for parents of 1, 2 or 3 children, single or in a single-earner couple.

The IMTR is the variation of disposable income per capita over a marginal change of pre-tax income. We compute the average IMTR over bins of 4% of the minimum wage to smooth the curve. The spikes we see are exit thresholds of various social transfers. After 60% minimum-wage, the unstable aspect of the implicit tax rate is linked to the calculation of housing benefit, in which resources are counted on an annual basis, rounded to the nearest hundred euros. The level of gain in employment therefore depends on whether the extra euros earned tip resources into the next highest hundred euros or not[67]. This poorly designed interaction between these two parameters generates very strange variations in marginal tax rates. For instance in these simulations, a single parent with two children going from 42% to 43% of the minimum wage undergoes an implicit marginal tax rate of 337% and these € 13 additional euros turn into an overall reduction of € 33 of disposable incomes. While the variations may be small in magnitude, they add a lot of noise to a system that is already hard to understand. For households, this means income uncertainty and unexplained variation, especially since these changes occur because of aggregated variations in income over the past 12 months.

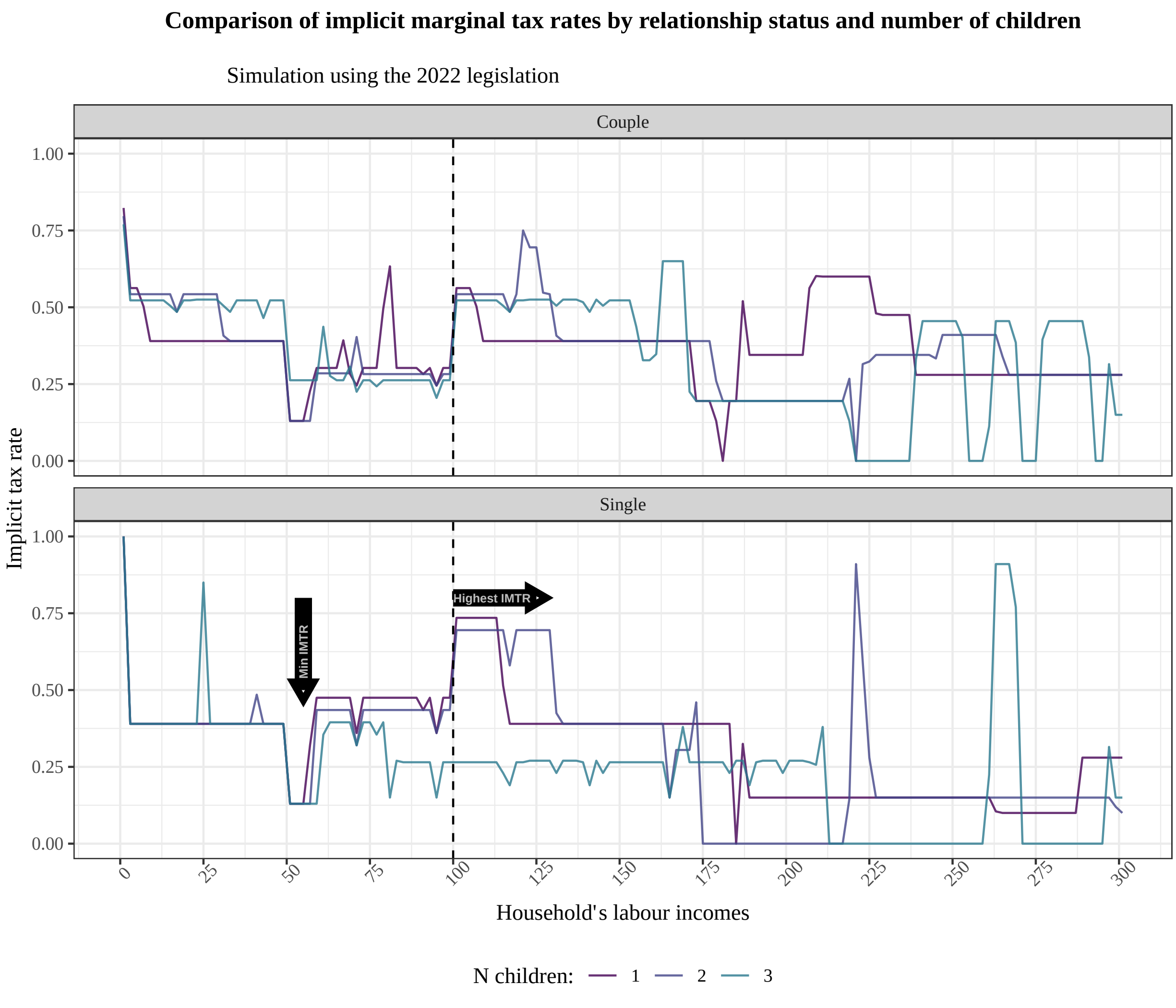
Apart from these small variations, there are four parts of the distribution we should comment. First, for households earning less than 50% minimum wage, the IMTR is higher four couples than single parents. Second, individual bonus of the PA starts at 50% minimum wage and creates a sharp drop in the IMTR up to 60%. Third, from this point, housing benefits start to decrease and, importantly, the IMTR is about 50% for single parents and roughly 38%

[67] See https://evaluation.securite-sociale.fr/home/financement/29-assurer-un-revenu-disponible.html

for couples, local variation aside. Fourth, the IMTR sharply increases at the full-time minimum wage for couples and single parents, but couples only reach 52% i.e. roughly the same level as for single parents from 60% to 100%. The latter face an IMTR of about 70% up to 115% to 130% of the minimum wage.

This brings us to the last important difference to notice. There are sharp differences by number of children. In particular, single parents of 3 children do not receive welfare or in-work benefits from 75% of the minimum wage. For them, social transfers depend mostly on housing benefits on the one hand, and family allowance on the other. Housing benefits increase with the number of children, and for parents of three with low or no income, there is an additional family allowance called *Complément familial*. All are entirely deduced from RSA and PA, explaining the little amounts and early exit.

Figure A.19: implicit marginal tax rate for single and couples parents by number of children



Sources: DREES, EDIFIS.
Note: The implicit marginal tax rate is directly computed in the simulation. We use an average over bins of 2pp of the minimum wage. Spikes with more than 100% marginal tax rate occuring at the exit thresholds of each transfers removed for clarity.

B Data preparation and cleaning

B.I Main databases: characteristics and pre-treatment

Description of the source files The source file contains three databases built from the `ALLSTAT FR6` monthly records from the National family allowance fund matched with design variables. All baseline table contains 175 columns.

- `BaselineClean` contains 2662 observations of the 5 cohorts, including those with unfavourable initial assessment and the 53 households from the pilot study. It was extracted from the complete dataset at the month before random assignment.
- `Reliance` contains 207636 individual $\times$ months observations, with missing values when unobserved
- `RelianceNoMissing` contains 194264 individual $\times$ months observations where lost households have been dropped.

We load and match two additional sources from INSEE: monthly minimum income and consumer price index for the bottom 20% of the national income distribution with reference value in 2015. Since many social transfers depend on the full-time minimum wage-level, it is useful to also define labour incomes as share of a full-time minimum wage. Then, we can use these variables with simulations of the tax-benefit system and use the *static* implicit marginal taxation rate in bunching estimates.

We drop the data of a file whose separation occurred the month of sampling from provisional datasets and whose status was only stabilised 2 months after random assignment. This household was in the encouragement group of the 2019 cohort and did not participate.

Pre-treatment of baseline characteristics There are 32 (1.5%) housholds whose initial incomes are missing, but whose total disposable incomes indicate that they have no income. We therefore impute 0 for these files. There are also 6 households who have no children in custody at that month. They all have one child with intermittent patterns of custody. We assign sample means for children age at baseline.

All monetary measures are converted in 2015 values, and we *winsorise* the .1% with extreme values, most likely due to misreports. We create dummies for the covariates used in the analysis and *distributional dummies* $i.e.$ for fractile of income distributions.

We centre all continuous covariates by the sample mean in their block. This transformation is useful when using models in long difference as it retrieves the average of the outcome at the observed month while removing baseline differences between blocks.

This baseline table is then matched to the full dataset by their Unique id.

Main databases The main database contains data from January, 2017 to ‘r June, 2023 for 2662 households, among which 548 were excluded from the experimental sample and 53 from the pilot. Removing them from the 207636 observations leaves us with 161694 observations.

Since we can only observe the first cohort from -13 months from random assignment and the fourth cohort up to 30 months after, we only keep observations of cohort 2018 to 2021 between these relative months from random assignment. This leaves us with a dataset of 103066 observations.

Last, two blocks with 5 have less than two observations due to the two files that we removed because they were ineligible. We remove their 317 observations.

The final database contains 102749 observations for 1666 households.

Post-treatment and aggregated databases We construct three databases with aggregated outputs or limited time frame.

- Average income from 18 to 30
- Total income from 18 to 30
- Binned worked-month × individuals

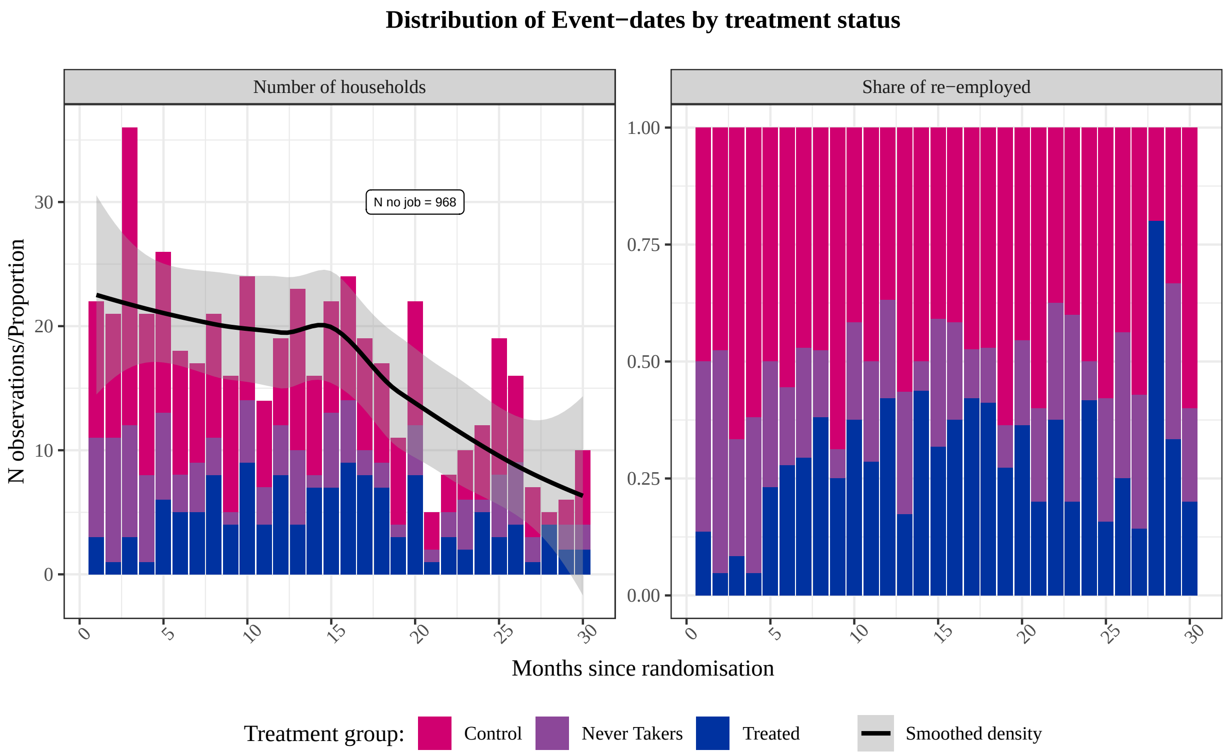
The two first datasets contain one observation per individual and compute the relevant aggregate outcomes. Some statistics are computed over the entire periods, others over months with positive labour incomes.

The last database records every wage in bins of 5% of the full-time minimum-wage over the same number of months for all households. This database contains as many different bins of income individuals ever recorded, and how many times they did. We use them for the bunching analysis.

C Additional descriptive statistics

C.I Distribution of dates of first job re-entry

Figure C.20: Compliers mostly start jobs in the last 6 months of the programme



Sources: ALLSTAT 2017−01−01 to 2023−06−01 cohorts 2018 to 2021 from 0 to 30 months from random assignment.
We omit the bar with 0 event for clarity and report their number in the label of the top panel.
Stacked bar chat of the number of households with first job re−entry at the month since random assignment.
Smoothed density use local polynomial regressions of the number of observation on group−time.

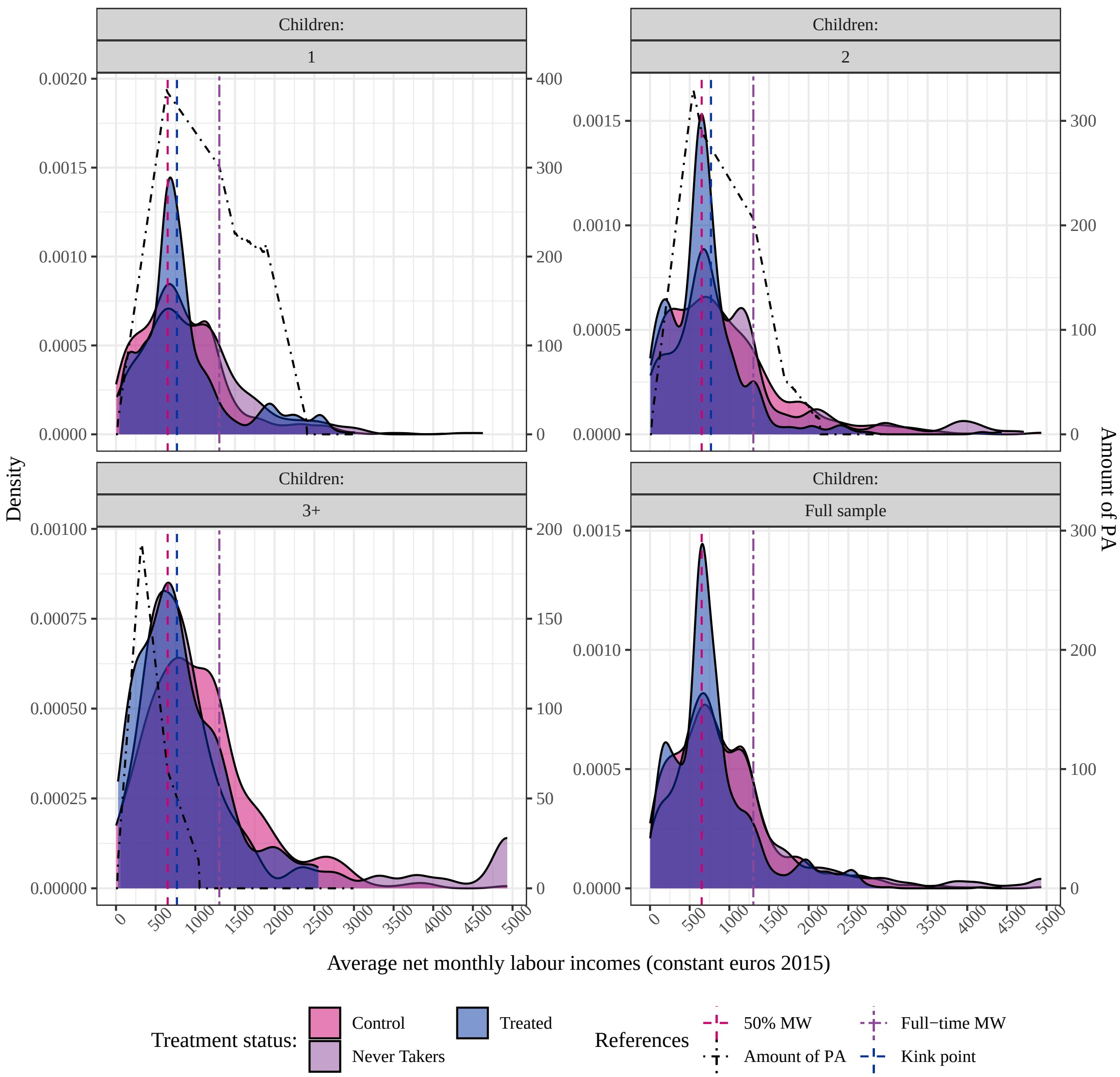
C.II Bunching at the individual labour income

C.III Mean differences in access to different sources of incomes

Figure C.21: Bunching of individual labour incomes at the sweet spot

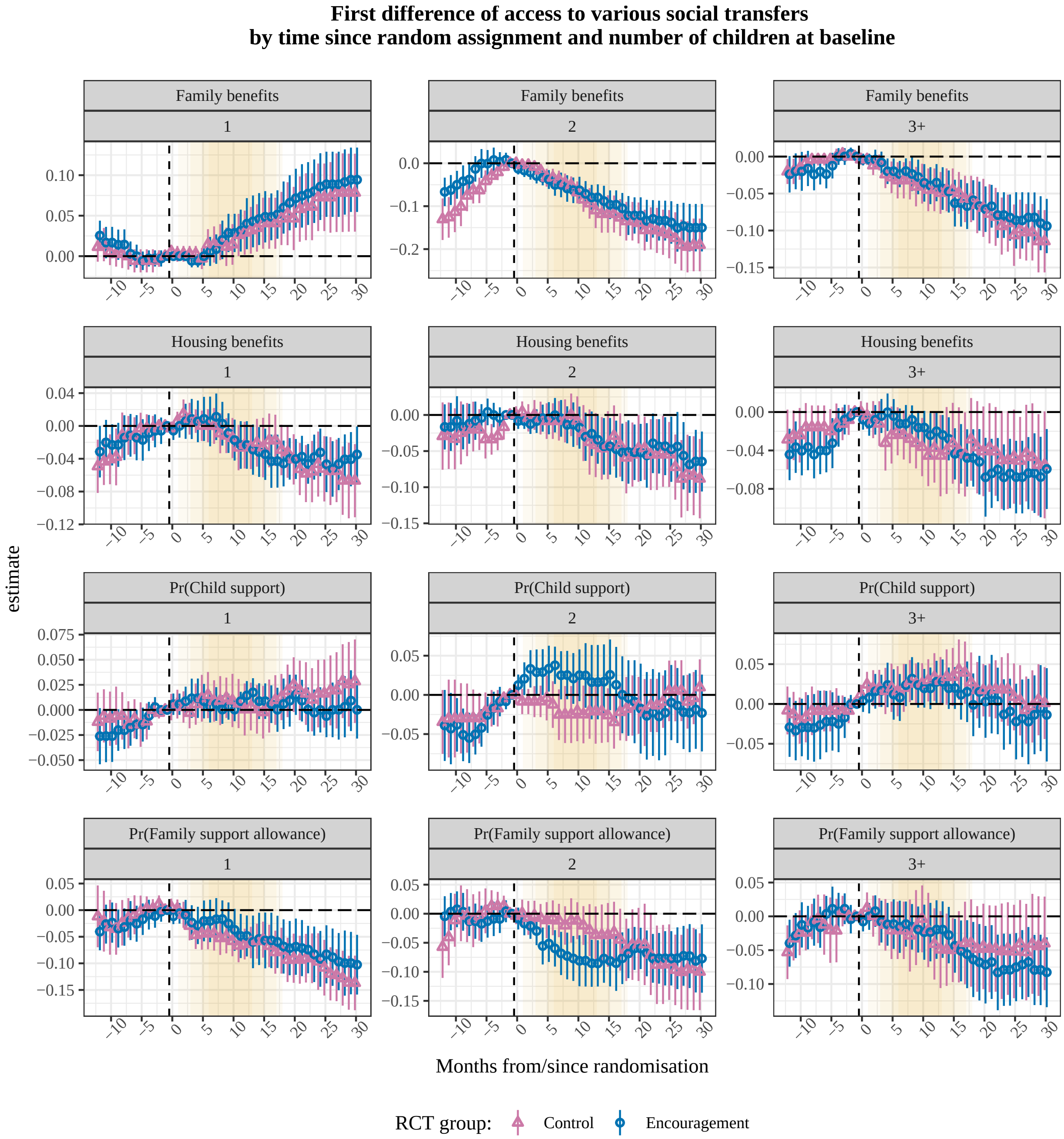
Distribution of individual labour incomes among those who work and theoretical amount of PA

Estimation over 12 months after the end of training



Sources: ALLSTAT, restricted sample over 12 months after the end of the programme among those who report positive labour incomes and smaller than 5 000 euros for clarity.
Notes: Kernel density of individual labour income for those with positive labour incomes.
The PA reference line indicates the theoretical amount of in−work benefits received for single parents by number of children and net labour income based on the EDIFIS model using the 2022 legislation.
Kink points indicate the level of income that minimises the implicit marginal tax rate.

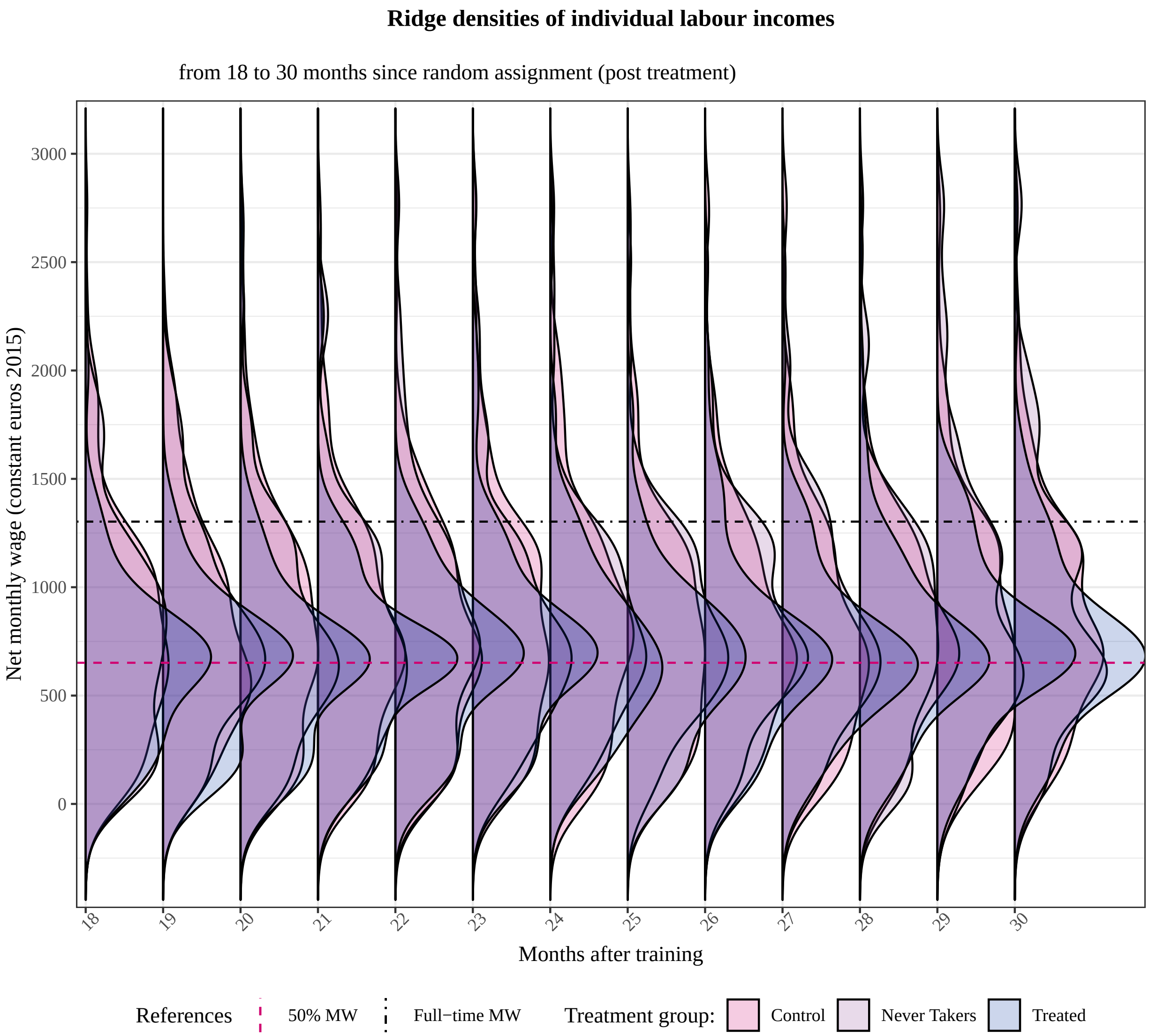
Figure C.22: Change in child support for mothers of 2, slower decrease of family benefits for mothers of 3 or more



Sources: ALLSTAT 2017−01−01 to 2023−06−01 cohorts 2018 to 2021 from −12 to 30 months from random assignment.
Means and point−wise 95% confidence intervals estimated by OLS regressions on month x group x encouragement dummies without intercept, using cluster−robust standard errors adjusted at the block x cohort level and inverse instrument propensity score weighting.
Rows correspond to different outcomes from separate regressions. Columns display results by number of children at baseline.

C.IV Monthly Densities of individual labour income from 18 to 30 months after random assignment

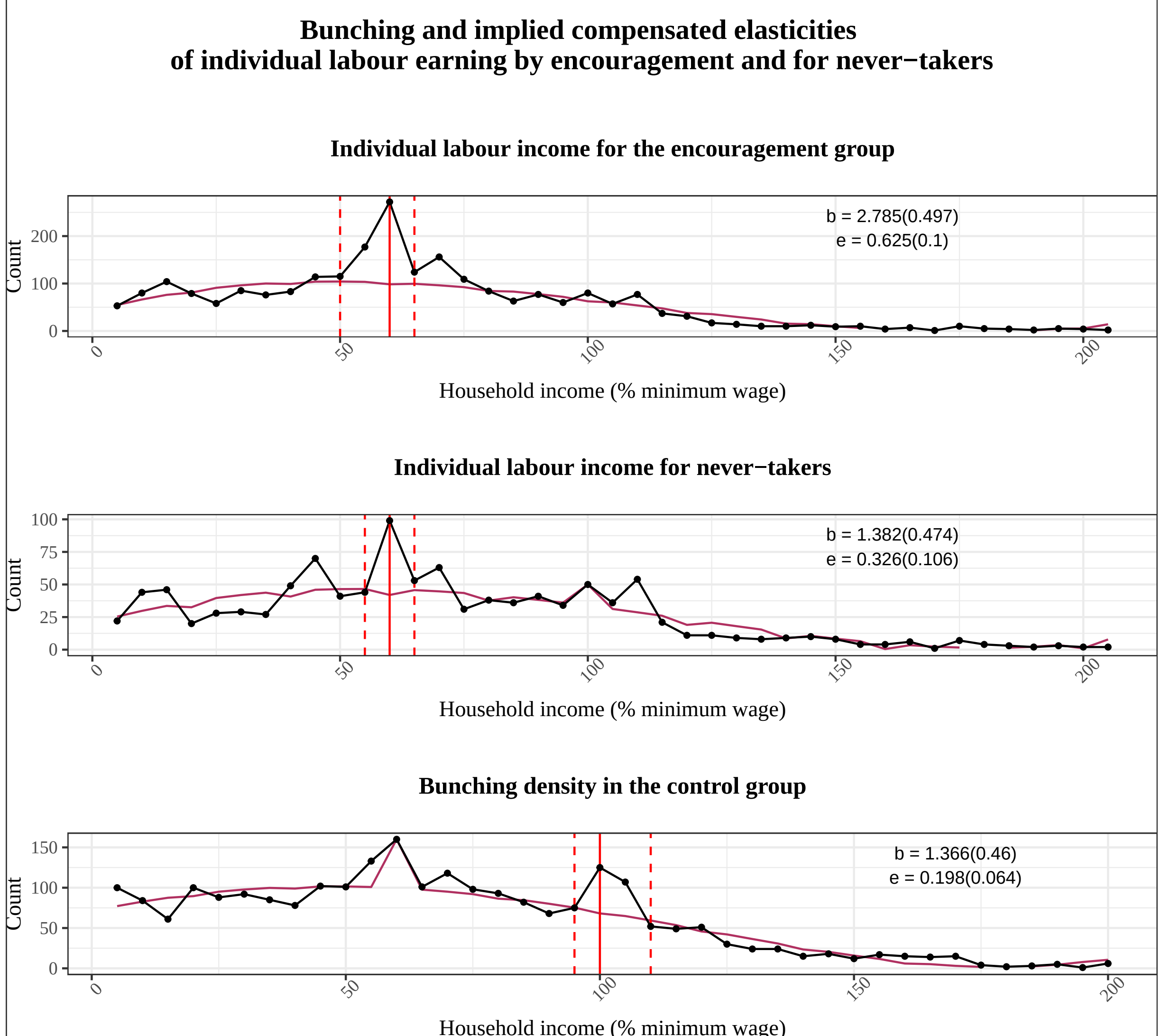
Figure C.23: Densities of individual labour incomes over the year after the end of training by number of children at baseline



Sources: ALLSTAT, cohorts 2018:2021, restricted sample over months 16 to 28 among those who report positive labour income.
Densities are scaled separately for each month of observation.

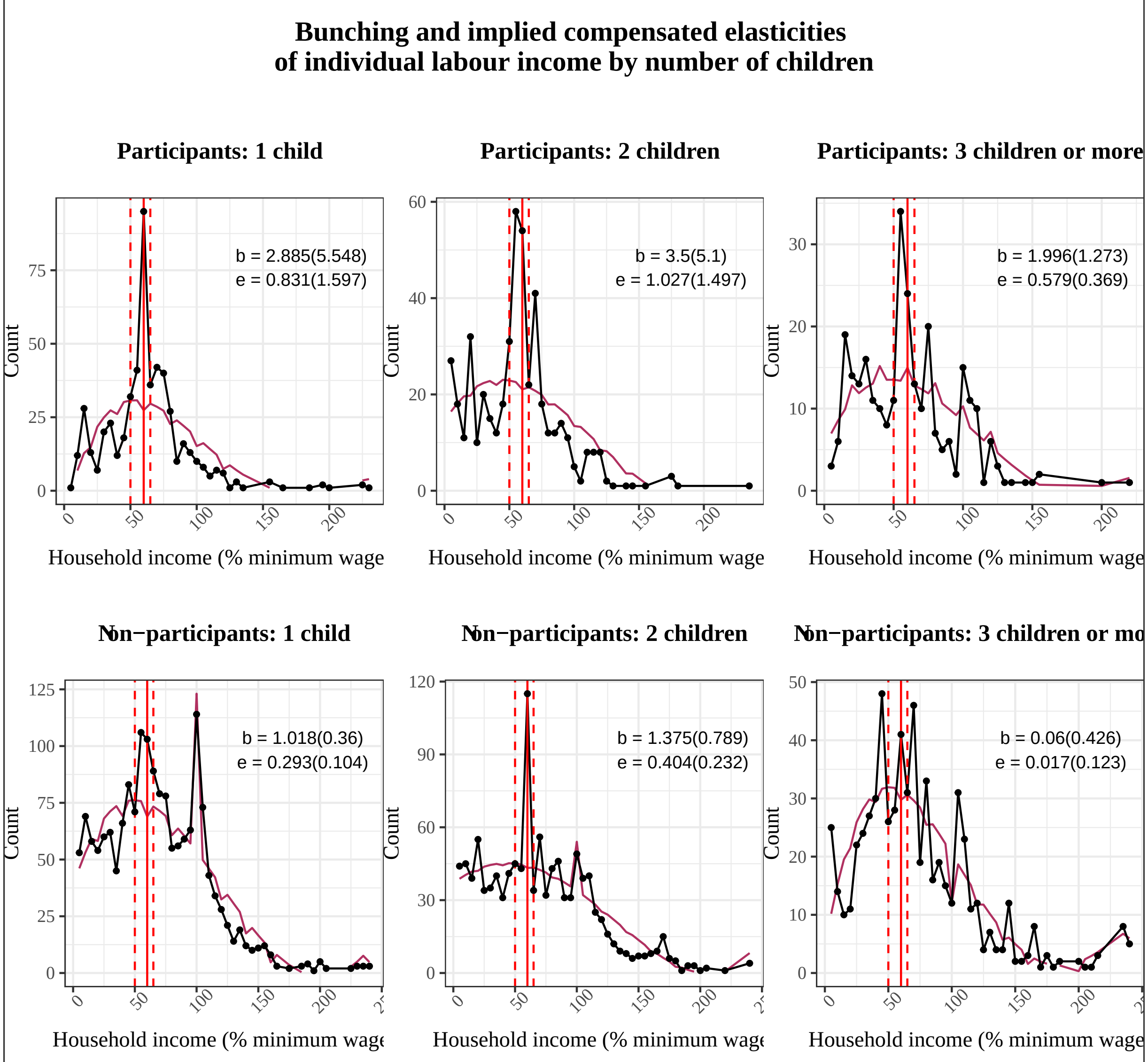
D Additional bunching estimates

Figure D.24: Comparing bunching of individual earnings between encouragement groups and never-takers



Sources: ALLSTAT 2017−01−01 to 2023−06−01 . Observations from 18 to 30 months after random assignment.
The database is collapsed into frequencies of individual labour incomes falling into bins of 5% minimum wage.
The statistical units are individual x months with positive labour incomes.
The top panel is from sub−sample of participants, the bottom panel is from the control group and never−takers.
Counterfactual densities estimated using polynomial of degree 4 , with round number dummies for every 20% SMIC corresponding to round number of days worked.
The bunching regions is delimited by the dashed red lines and we use the average
IMTR for single parents with 1,2,3 chidren presented in the table at the begining of the section.
Elasticities use parametric restriction and standard errors are computed by bootstrap using 499 replications.

Figure D.25: Bunching of individual incomes by number of children at baseline



Sources: ALLSTAT 2017−01−01 to 2023−06−01 . Observations from 18 to 30 months after random assignment.
The database is collapsed into frequencies of individual labour incomes falling into bins of 5% minimum wage.
The statistical units are individual x months with positive labour incomes.
The top panel is from sub−sample of participants, the bottom panel is from the control group and never−takers.
Counterfactual densities estimated using polynomial of degree 4 , with round number dummies for every 20% SMIC corresponding to round number of days worked.
The bunching regions are delimited by the dashed red lines and we use the average IMTR for single parents presented in the table in the Bunching section for elasticities.
Elasticities use parametric restriction and standard errors are computed by bootstrap using 499 replications.

E Additional estimations of counterfactual densities

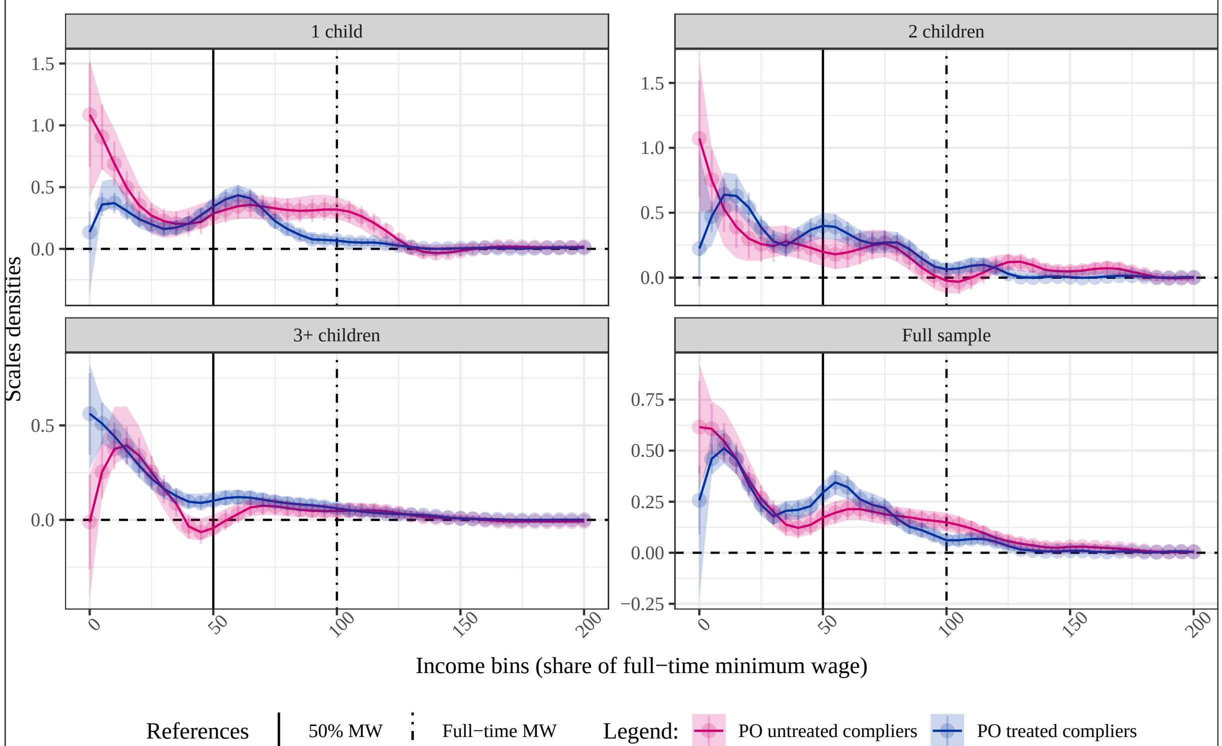
E.I Counterfactual densities by individual pre-tax incomes

Figure E.26: Counterfactual densities of compliers' individual income

Distribution of potential individual's pre−tax income of treated and untreated compliers

Non−parametric data driven estimation of compliers' counterfactual densities

Weighted local regression distribution estimates, Kappa weighting

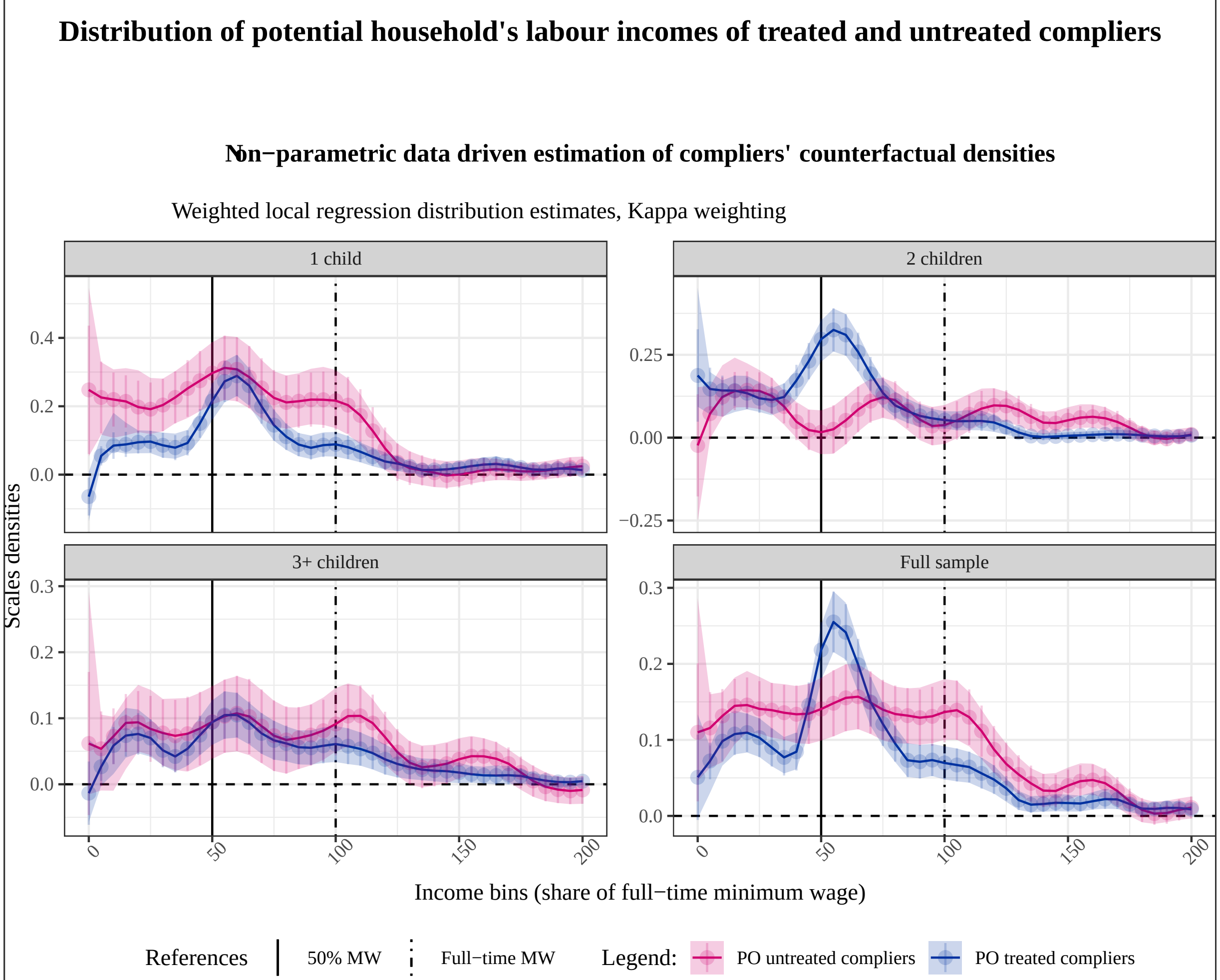


Sources: ALLSTAT, observations from 18 to 30 months since random assignment.
The dependent variable is the individual's pre−tax income as share of full−time minimum wage.
Notes: Estimations of the potential density functions for compliers following Cattaneo et al. (2021), using the R package 'lpdensity'.
Bandwidth chosen using regularised integrated MSE−optimal and the Epanechnikov kernel function.
Instrument propensity scores estimated by probit of encouragement on block fixed effects to smooth support.
Densities use local polynomial regressions of order 3 , simultaneous 95% confidence intervals use polynomial of order 4 , 1 degree higher for biais correction.
Densities are scaled over the support of potential income with positive values. The counterfactual probability of reporting 0 income is estimated by TSLS with block fixed effects using T x 1(Y<.000001) as outcome and d instrumented by the re−centred instrument and block fixed effect, with T = D for Y(1) and T=(1−D) for Y(0).

Weights are computed for each sub−group separately: $\kappa_0 = \frac{1}{1-D_i}\frac{1-Z_i-(1-\hat{q}_b)}{\hat{q}_b(1-\hat{q}_b)} \quad \kappa_1 = \frac{1}{D_i}\frac{Z_i-\hat{q}_b}{\hat{q}_b(1-\hat{q}_b)}$

E.II Counterfactual densities of household's incomes

Figure E.27: Counterfactual densities of compliers' household labour incomes



Sources: ALLSTAT, observations from 18 to 30 months since random assignment.
The dependent variable is the household's labour income as share of full−time minimum wage.
Notes: Estimations of the potential density functions for compliers following Cattaneo et al. (2021), using the R package 'lpdensity'.
Bandwidth chosen using regularised integrated MSE−optimal and the Epanechnikov kernel function.
Instrument propensity scores estimated by probit of encouragement on block fixed effects to smooth support.
Densities use local polynomial regressions of order 3 , simultaneous 95% confidence intervals use polynomial of order 4 , 1 degree higher for biais correction.
Densities are scaled over the support of potential income with positive values. The counterfactual probability of reporting 0 income is estimated by TSLS with block fixed effects using T x 1(Y<.000001) as outcome and d instrumented by the re−centred instrument and block fixed effect, with T = D for Y(1) and T=(1−D) for Y(0).

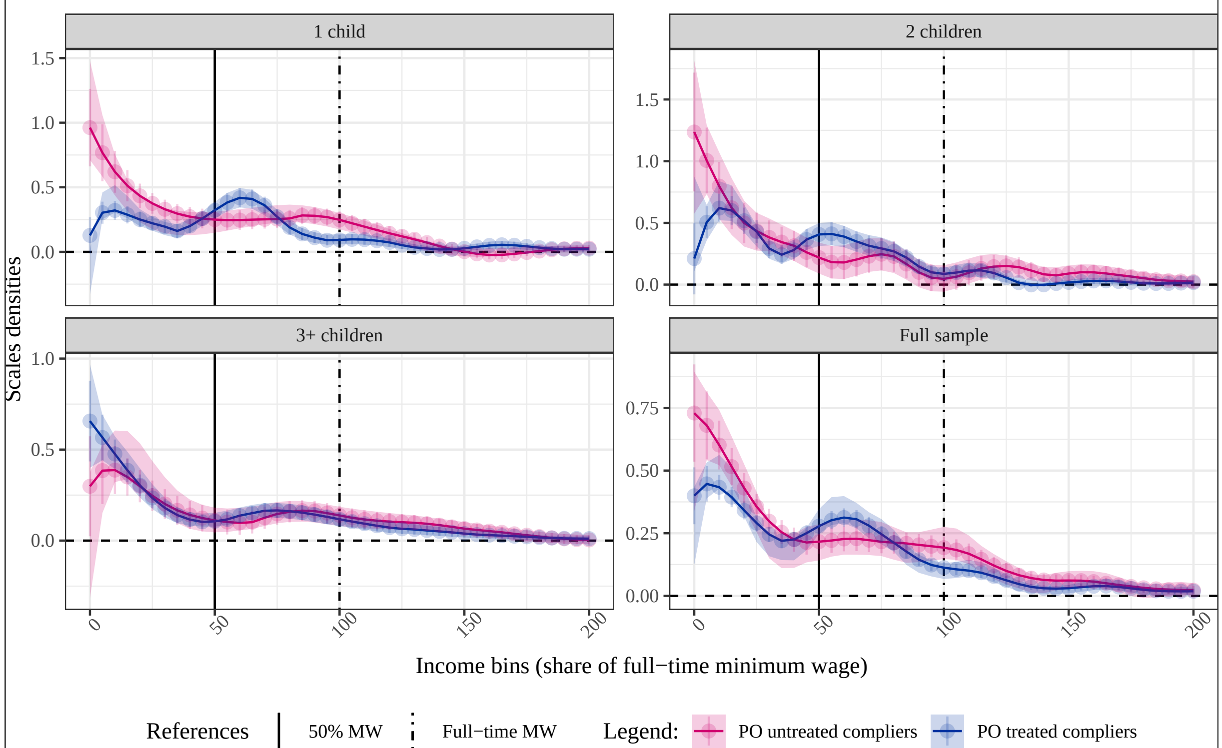
Weights are computed for each sub−group separately: $\kappa_0 = \frac{1}{1-D_i}\frac{1-Z_i-(1-\hat{q}_b)}{\hat{q}_b(1-\hat{q}_b)}$ $\kappa_1 = \frac{1}{D_i}\frac{Z_i-\hat{q}_b}{\hat{q}_b(1-\hat{q}_b)}$

Figure E.28: Counterfactual densities of compliers' household pre-tax income

Distribution of potential household's pre−tax income of treated and untreated compliers

Non−parametric data driven estimation of compliers' counterfactual densities

Weighted local regression distribution estimates, Kappa weighting

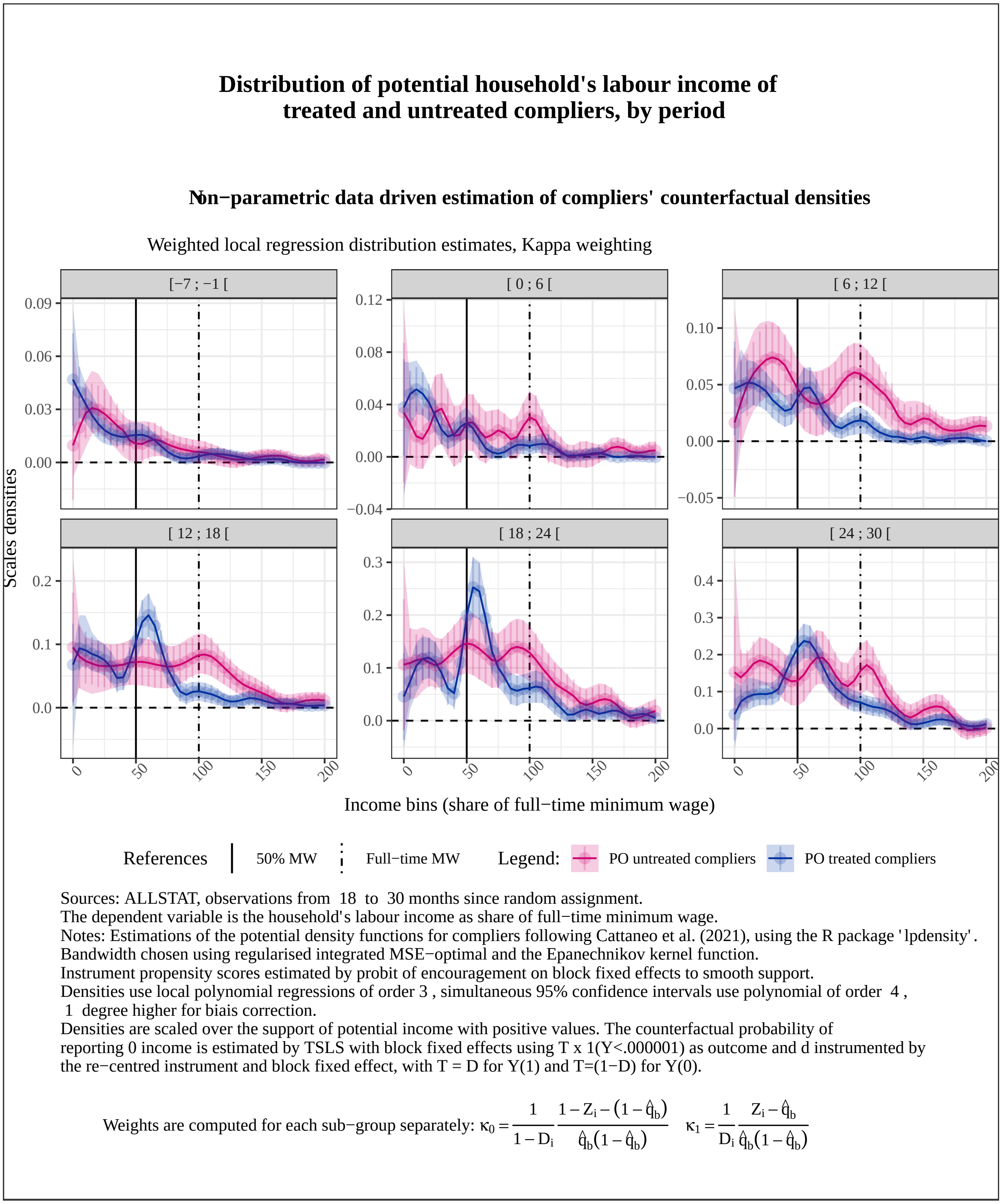


Sources: ALLSTAT, observations from 18 to 30 months since random assignment.
The dependent variable is the household's pre−tax income as share of full−time minimum wage.
Notes: Estimations of the potential density functions for compliers following Cattaneo et al. (2021), using the R package 'lpdensity'.
Bandwidth chosen using regularised integrated MSE−optimal and the Epanechnikov kernel function.
Instrument propensity scores estimated by probit of encouragement on block fixed effects to smooth support.
Densities use local polynomial regressions of order 3, simultaneous 95% confidence intervals use polynomial of order 4, 1 degree higher for biais correction.
Densities are scaled over the support of potential income with positive values. The counterfactual probability of reporting 0 income is estimated by TSLS with block fixed effects using T x 1(Y<.01) as outcome and d instrumented by the re−centred instrument and block fixed effect, with T = D for Y(1) and T=(1−D) for Y(0).

Weights are computed for each sub−group separately: $\kappa_0 = \frac{1}{1-D_i}\frac{1-Z_i-(1-\hat{q}_b)}{\hat{q}_b(1-\hat{q}_b)}$ $\kappa_1 = \frac{1}{D_i}\frac{Z_i-\hat{q}_b}{\hat{q}_b(1-\hat{q}_b)}$

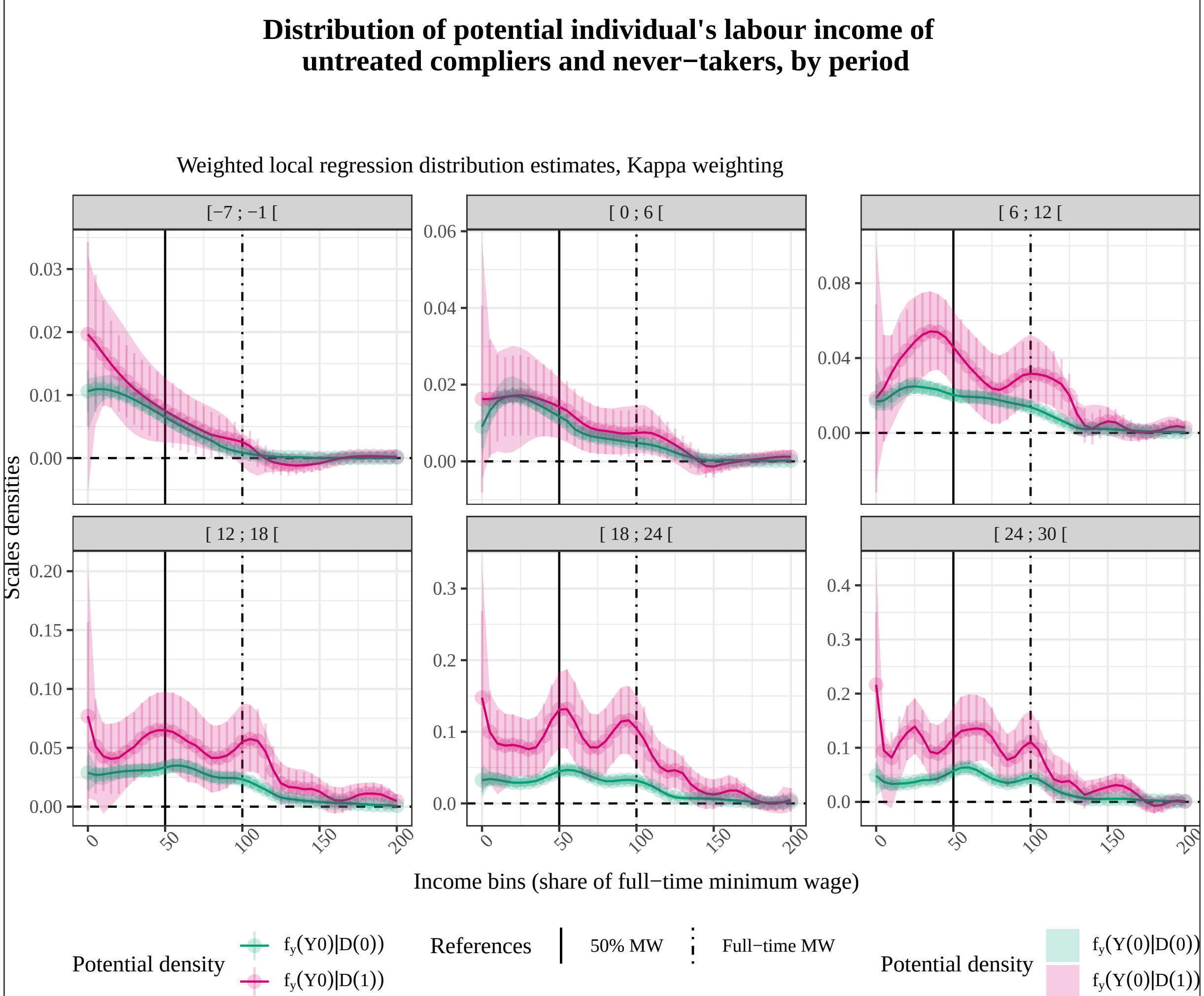
E.III Evolution of bunching over 6 month periods

Figure E.29: Evolution of potential household's labour income by 6 month periods



E.IV Comparisons of individual labour incomes with never-takers

Figure E.30: Comparing potential individual labour incomes of untreated compliers and never-takers

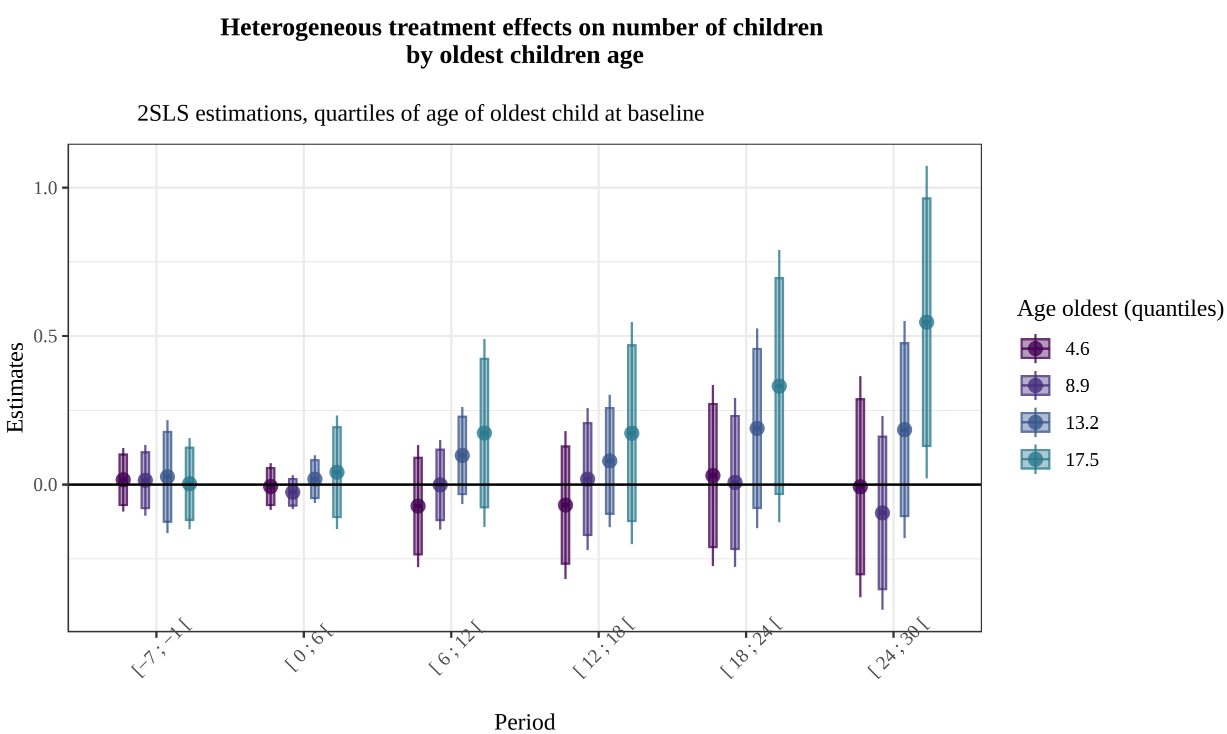


Sources: ALLSTAT, observations from 18 to 30 months since random assignment.
Notes: Estimations of the potential density functions following Cattaneo et al. (2021), using the R package 'lpdensity'.
Bandwidth chosen using regularised integrated MSE−optimal and the Epanechnikov kernel function.
Instrument propensity scores estimated by probit of encouragement on block fixed effects to smooth support.
Densities use local polynomial regressions of order 2 , simultaneous confidence intervals use polynomial of order 3 , 1 higher to correct bias.
Densities are scaled over the support of potential income with positive values. The counterfactual probability of reporting 0 income is estimated by TSLS with block fixed effects using T x 1(Y<.00001) as outcome and d instrumented by the re−centred instrument and block fixed effect, T=(1−D) for Y(0).

Weights for untreated compliers: $\kappa_0 = \frac{1}{1-D_i}\frac{1-Z_i-(1-\hat{q}_b)}{\hat{q}_b(1-\hat{q}_b)}$

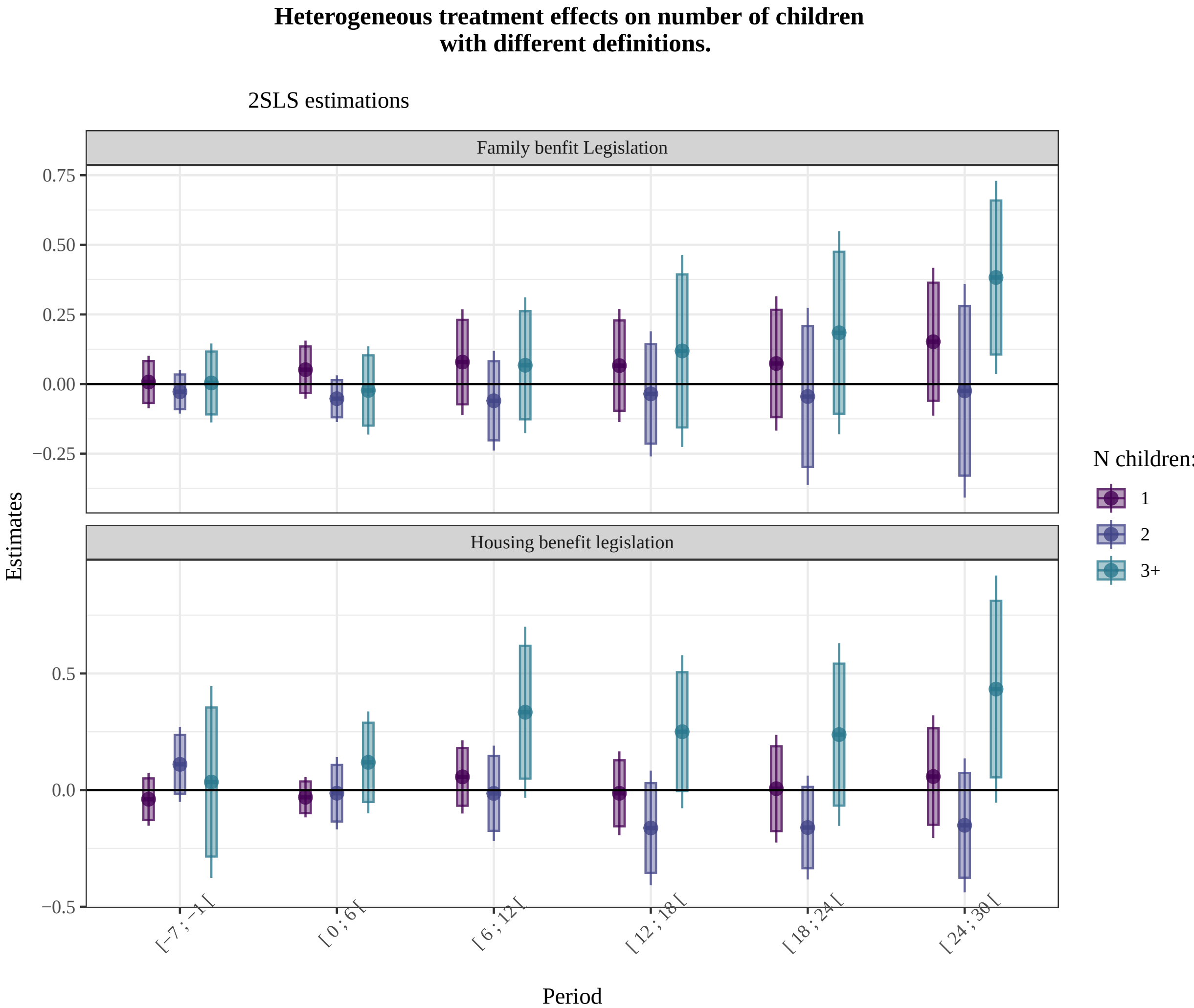
F Other estimations on family structure

Figure F.31: Heterogenous treatment effects on number of children, by quartile of oldest children at baseline



Sources: ALLSTAT 2017−01−01 to 2023−06−01 cohorts 2018 to 2021.
Notes: The dependent variable is the number of children under responsibility as used for family allowance benefits.
Cluster−robust standard errors at the block x cohort level.
− Error bars indicate 95 % pointwise confidence intervals and extended lines account for FWER by subgroup.
− All models include blocks x cohort x relative months fixed effects and covariates listed in the paper.

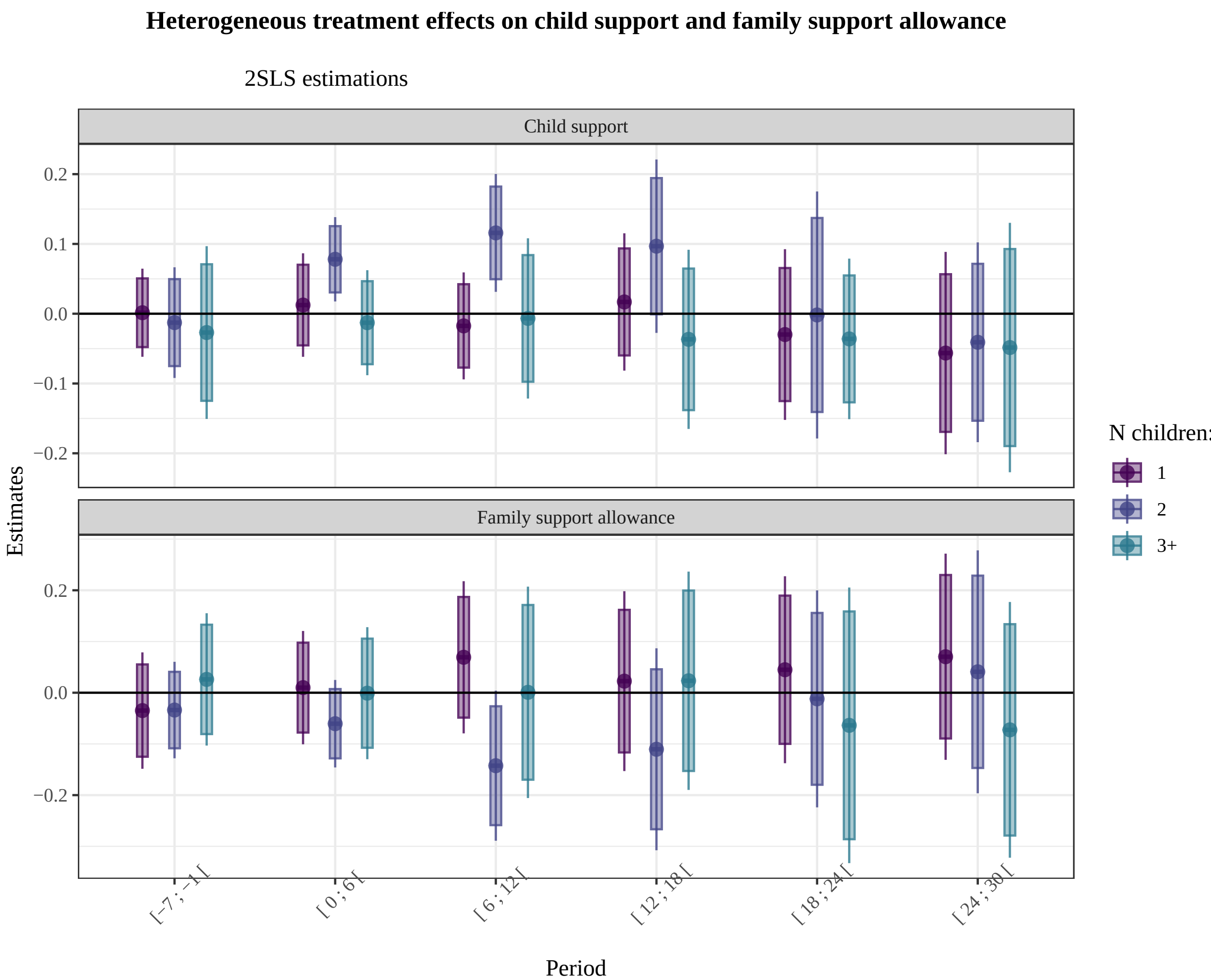
Figure F.32: Effect on different definition of number of children



Sources: ALLSTAT 2017−01−01 to 2023−06−01 cohorts 2018 to 2021.
Notes: The dependent variable is the number of dependent children either defined by familiy benefits or housing benefits.
Cluster−robust standard errors at the block x cohort level.
− Error bars indicate 95 % pointwise confidence intervals and extended lines account for FWER by subgroup.
− All models include blocks x cohort x relative months fixed effects and instrument propensity score weighting.

F.I Differential effects on child support and family support allowance

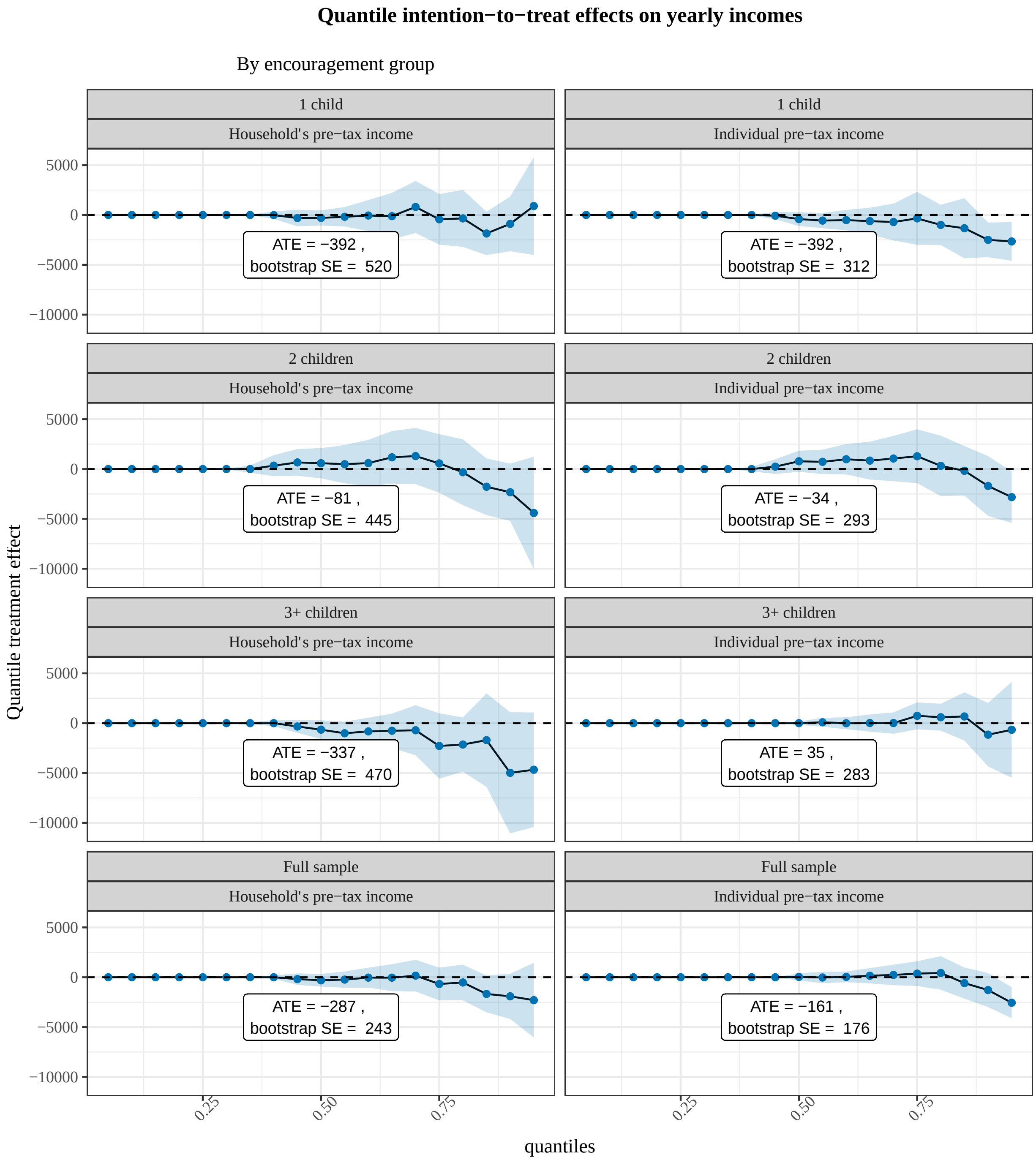
Figure F.33: Perfect substition of ASF and child support for parents of 2



Sources: ALLSTAT 2017−01−01 to 2023−06−01 cohorts 2018 to 2021.
Notes: The dependent variables are long difference between the date and the month before random assignment.
Child support is paid by non−custodial parents and deduced from other social transfers.
Family support allowance is paid in substitute of child support.
Cluster−robust standard errors at the block x cohort level.
− Error bars indicate 95 % pointwise confidence intervals and extended lines account for FWER by subgroup.
− All models include blocks x cohort x relative months fixed effects and instrument propensity score weighting.

F.II Quantile treatment effects on yearly pre-tax incomes

Figure F.34: Quantile treatment effects on cumulative pre-tax incomes over the year after the end of the programme



Sources: ALLSTAT, observations from 18 to 30 months since random assignment.
Notes: Estimations of the quantile intention−to−treat effect controling for blocks x cohort by inverse−propensity score weightingfollowing Firpo (2007). 95% Confidence intervals estimated by bootstrap.

Bibliography

Abadia, Laura, Ilf Bencheikh, Marie Borel, and Marc Gurgand. 2017. "Accompagnement Des Chômeurs de Longue Durée: Revue Systématique Des Évaluation d'impact." Paris: J-pal.
Abadie, A. 2003. "Semiparametric Instrumental Variable Estimation of Treatment Response Models." *Journal of Econometrics* 113 (2): 231–63.
Abadie, Alberto. 2005. "Semiparametric Difference-in-Differences Estimators." *The Review of Economic Studies* 72 (1): 1–19.
Abadie, Alberto, Joshua Angrist, and Guido Imbens. 2002. "Instrumental Variables Estimates of the Effect of Subsidized Training on the Quantiles of Trainee Earnings." *Econometrica* 70 (1): 91–117.
Abadie, Alberto, Susan Athey, Guido W. Imbens, and Jeffrey M. Wooldridge. 2020. "Sampling□Based Versus Design□Based Uncertainty in Regression Analysis." *Econometrica* 88 (1): 265–96.
Abadie, Alberto, Susan Athey, Guido W Imbens, and Jeffrey M Wooldridge. 2022. "When Should You Adjust Standard Errors for Clustering?" *The Quarterly Journal of Economics* 138 (1): 1–35.
Abbas, Hicham, and Bertrand Garbinti. 2019. "De La Rupture Conjugale à Une Éventuelle Remise En Couple : L'évolution Des Niveaux de Vie Des Familles Monoparentales Entre 2010 Et 2015." *INSEE Références*, France portrait social,.
Abdulkadiroglu, Atila, Joshua D Angrist, Yusuke Narita, and Parag A Pathak. 2017. "Research Design Meets Market Design: Using Centralized Assignment for Impact Evaluation." *Econometrica* 85 (5): 1373–1432.
Abdulkadiroğlu, Atila, Parag A. Pathak, and Alvin E. Roth. 2009. "Strategy-Proofness Versus Efficiency in Matching with Indifferences: Redesigning the NYC High School Match." *The American Economic Review* 99 (5): 1954–78.
Abdulkadiroğlu, Atila, Parag A. Pathak, and Christopher R. Walters. 2018. "Free to Choose: Can School Choice Reduce Student Achievement?" *American Economic Journal: Applied Economics* 10 (1): 175–206.
Abdulkadiroğlu, Atila, and Tayfun Sönmez. 2003. "School Choice: A Mechanism Design Approach." *American Economic Review* 93 (3): 729–47.
———. 2013. "Matching Markets: Theory and Practice." In *Advances in Economics and Econometrics*, edited by Daron Acemoglu, Manuel Arellano, and Eddie Dekel, 3–47. Cambridge: Cambridge University Press.
Abrassart, Aurélien, and Giuliano Bonoli. 2015. "Availability, Cost or Culture? Obstacles to Childcare Services for Low-Income Families." *Journal of Social Policy* 44 (4): 787–806.
Acemoglu, Daron, and David Autor. 2012. "What Does Human Capital Do? A Review of Goldin and Katz's "The Race Between Education and Technology"." *Journal of Economic Literature* 50 (2): 426–63.
Acs, Marie, Bertrand Lhommeau, and Emilie Raynaud. 2015. "Les Familles Monoparentales Depuis 1990." *Drees*, Solidarité et santé, no. 67: 36.
Adam, Stuart, Antoine Bozio, Carl Emmerson, David Greenberg, and Genevieve Knight. 2008. "A Cost-Benefit Analysis of Pathways to Work for New and Repeat Incapacity Benefits Claimants." 498. Department for Work and Pensions, Her Majesty's stationery office.
Akbarpour, Mohammad, Julien Combe, Yinghua He, Victor Hiller, Robert Shimer, and Olivier Tercieux. 2020. "UNPAIRED KIDNEY EXCHANGE: OVERCOMING DOUBLE COINCIDENCE OF WANTS WITHOUT MONEY." *NBER Working Papers*.
Akbarpour, Mohammad, Adam Kapor, Christopher Neilson, Winnie van Dijk, and Seth Zimmerman. 2022. "Centralized School Choice with Unequal Outside Options." *Journal of Public Economics* 210 (June): 104644.
Akerlof, George A., Janet L. Yellen, and Michael L. Katz. 1996. "An Analysis of Out-of-Wedlock Childbearing in the United States*." *The Quarterly Journal of Economics* 111 (2): 277–317.
Albelda, Randy, Susan Himmelweit, and Jane Humphries. 2004. "The Dilemmas of Lone Motherhood: Key Issues for Feminist Economics." *Feminist Economics* 10 (2): 1–7.
Albright, Alex. 2023. "The Hidden Effects of Algorithmic Recommendations," July.
Alesina, Alberto, Andrea Ichino, and Loukas Karabarbounis. 2011. "Gender-Based Taxation and the Division of Family Chores." *American Economic Journal: Economic Policy* 3 (2): 1–40.
Alesina, Alberto, Stefanie Stantcheva, and Edoardo Teso. 2018. "Intergenerational Mobility and Preferences for Redistribution." *The American Economic Review* 108 (2): 521–54.
Allègre, Guillaume. 2011. "Le RSA : Redistribution Vers Les Travailleurs Pauvres Et Offre de Travail." *Revue de l'OFCE* 118 (3): 33–61.
———. 2024. "Les Nouvelles Lois Sur Les Pauvres (1989-2023)."
Allègre, Guillaume, Hélène Périvier, and Muriel Pucci. 2021. "Taxation of Couples and Marital Status – Simulation of Three Reforms of the Marital Quotient in France." *Economie Et Statistique / Economics and Statistics*, no.

526-527 (October): 3–20.
Allègre, Guillaume, Véronique Simonnet, and Catherine Sofer. 2015. "Child Care and Labour Market Participation in France: Do Monetary Incentives Matter?" *Annals of Economics and Statistics*, no. 117/118: 115–39.
Allen, Jim, and Rolf van der Velden. 2001. "Educational Mismatches Versus Skill Mismatches: Effects on Wages, Job Satisfaction, and on the job Search." *Oxford Economic Papers* 53 (3): 434–52.
Ananat, Elizabeth Oltmans, Jonathan Gruber, Phillip B Levine, and Douglas Staiger. 2009. "Abortion and Selection." *The Review of Economics and Statistics* 91 (1): 124–36.
Ananat, Elizabeth Oltmans, and Daniel M. Hungerman. 2012. "The Power of the Pill for the Next Generation: Oral Contraception's Effects on Fertility, Abortion, and Maternal and Child Characteristics." *The Review of Economics and Statistics* 94 (1): 37–51.
Ananat, Elizabeth O., and Guy Michaels. 2008. "The Effect of Marital Breakup on the Income Distribution of Women with Children." *The Journal of Human Resources* 43 (3): 611–29.
Anderson, Nina J., Susan A. Graham, Heather Prime, Jennifer M. Jenkins, and Sheri Madigan. 2021. "Linking Quality and Quantity of Parental Linguistic Input to Child Language Skills: A Meta-Analysis." *Child Development* 92 (2): 484–501.
Andresen, Martin Eckhoff, and Emily Nix. 2022. "What Causes the Child Penalty? Evidence from Adopting and Same-Sex Couples." *Journal of Labor Economics*, October.
Angrist, J. D., and Guido W Imbens. 1995. "Two-Stage Least Squares Estimation of Average Causal Effects in Models With Variable Treatment Intensity." *Journal of the American Statistical Association* 90 (430).
Angrist, Joshua D. 2022. "Empirical Strategies in Economics: Illuminating the Path From Cause to Effect." *Econometrica* 90 (6): 2509–39.
Angrist, Joshua D., Peter D. Hull, Parag A. Pathak, and Christopher R. Walters. 2017. "Leveraging Lotteries for School Value-Added: Testing and Estimation*." *The Quarterly Journal of Economics* 132 (2): 871–919.
Angrist, Joshua D, Guido W Imbens, and Donald B Rubin. 1996. "Identification of Causal Effects Using Instrumental Variables." *Journal of the American Statistical Association* 91 (434): 444–55.
Angrist, Joshua D, and Jörn-Steffen Pischke. 2010. "The Credibility Revolution in Empirical Economics: How Better Research Design Is Taking the Con Out of Econometrics." *Journal of Economic Perspectives* 24 (2): 3–30.
Angrist, Joshua, and William Evans. 1998. "Children and Their Parents' Labor Supply: Evidence from Exogenous Variation in Family Size." *The American Economic Review , Jun., 1998, Vol. 88, No. 3*, June, w5778.
Angrist, Joshua, Peter Hull, and Christopher Walters. 2023. "Methods for Measuring School Effectiveness." In *Handbook of the Economics of Education*. Elsevier.
Angrist, Joshua, and Michal Kolesár. 2022. "One Instrument to Rule Them All: The Bias and Coverage of Just-ID IV." December 15, 2022.
Anne, Denis, Laetitia Challe, Pascale Petit, Sylvain Chareyron, Yannick L'Horty, and Loïc du Parquet. 2019. "Discrimination Dans Le Recrutement Des Grandes Entreprises : Une Approche Multicanal." Rapport d'évaluation. Paris: CGET.
Appelbaum, Lauren D. 2001. "The Influence of Perceived Deservingness on Policy Decisions Regarding Aid to the Poor." *Political Psychology* 22 (3): 419–42.
Arni, Patrick, Rafael Lalive, and Jan C. Van Ours. 2013. "How Effective Are Unemployment Benefit Sanctions? Looking Beyond Unemployment Exit." *Journal of Applied Econometrics* 28 (7): 1153–78.
Ashworth, Karl, Andreas Cebulla, David Greenberg, and Robert Walker. 2004. "Meta-Evaluation: Discovering What Works Best in Welfare Provision." *Evaluation* 10 (2): 193–216.
Askenazy, Philippe. 2022. "The Transformations of the French Labor Market, 2000–2021." *IZA World of Labor*.
Athari, Elika. 2023. "Deux Tiers Des Bénéficiaires Du RSA Au Chômage Se Déclarent Freinés Dans Leurs Démarches de Recherche d'emploi." Etudes et résultats 1252. Paris: DREES.
Athey, S., and G. W. Imbens. 2017a. "Chapter 3 - the Econometrics of Randomized Experiments." In *Handbook of Field Experiments*, edited by Abhijit Vinayak Banerjee and Esther Duflo, 1:73–140. Handbook of Economic Field Experiments. North-Holland.
———. 2017b. "Chapter 3 - The Econometrics of Randomized Experiments." In *Handbook of Economic Field Experiments*, edited by Abhijit Vinayak Banerjee and Esther Duflo, 1:73–140. Handbook of Field Experiments. North-Holland.
Athey, Susan, and Guido W. Imbens. 2017. "The State of Applied Econometrics: Causality and Policy Evaluation." *Journal of Economic Perspectives* 31 (2): 3–32.

Attanasio, Orazio. 2022. “Public Childcare, Labor Market Outcomes of Caregivers, and Child Development: Experimental Evidence from Brazil,” 55.

Attanasio, Orazio P. 2015. “The Determinants of Human Capital Formation During the Early Years of Life: Theory, Measurement, and Policies.” *Journal of the European Economic Association* 13 (6): 949–97.

Attanasio, Orazio, Peter Levell, Hamish Low, and Virginia Sánchez-Marcos. 2018. “Aggregating Elasticities: Intensive and Extensive Margins of Women’s Labor Supply.” *Econometrica* 86 (6): 2049–82.

Autor, David H, Susan N Houseman, and Sari Pekkala Kerr. 2017. “The Effect of Work First Job Placements on the Distribution of Earnings: An Instrumental Variable Quantile Regression Approach.” *Journal of Labor Economics* 35 (1): 42.

Autor, David, Claudia Goldin, and Lawrence F. Katz. 2020. “Extending the Race Between Education and Technology.” *AEA Papers and Proceedings* 110 (May): 347–51.

Avenel, Cyprien, Marine Boisson Cohen, Sandrine Dauphin, Nicolas Duvoux, Christophe Fourel, Manon Julien, and Bruno Palier. 2017. *L’investissement Social : Quelle Stratégie Pour La France ?* La Documentation Française. Paris.

Avery, Christopher, and Parag A. Pathak. 2021. “The Distributional Consequences of Public School Choice.” *American Economic Review* 111 (1): 129–52.

Avram, Silvia, Mike Brewer, and Andrea Salvatori. 2018. “Can’t Work or Won’t Work: Quasi-experimental Evidence on Work Search Requirements for Single Parents.” *Labour Economics* 51 (April): 63–85.

Badgett, M. V. Lee, Christopher S. Carpenter, and Dario Sansone. 2021. “LGBTQ Economics.” *Journal of Economic Perspectives* 35 (2): 141–70.

Bailey, Drew, Greg J. Duncan, Candice L. Odgers, and Winnie Yu. 2017. “Persistence and Fadeout in the Impacts of Child and Adolescent Interventions.” *Journal of Research on Educational Effectiveness* 10 (1): 7–39.

Banerjee, Abhijit, Esther Duflo, and Garima Sharma. decembre 2021. “Long-Term Effects of the Targeting the Ultra Poor Program.” *American Economic Review: Insights* 4 (3): 471–86.

Barbuscia, Anna, Ariane Pailhé, Anne Solaz, Nathalie Bajos, Josiane Warszawski, Guillaume Bagein, François Beck, et al. 2023. “Do Income and Employment Uncertainty Affect Couple Stability? Evidence for France During the COVID-19 Pandemic.” *European Journal of Population* 39 (1): 19.

Bargain, Olivier, Mathias Dolls, Dirk Neumann, Andreas Peichl, and Sebastian Siegloch. 2014. “Tax-Benefit Revealed Social Preferences in Europe and the US.” *Annals of Economics and Statistics*, no. 113/114: 257–89.

Bargain, Olivier, and Karina Doorley. 2011. “Caught in the Trap? Welfare’s Disincentive and the Labor Supply of Single Men.” *Journal of Public Economics* 95 (9-10): 1096–1110.

Bargain, Olivier, and Augustin Vicard. 2014. “Le RMI Et Son Successeur Le RSA Découragent-Ils Certains Jeunes de Travailler ? Une Analyse Sur Les Jeunes Autour de 25 Ans.” *Economie Et Statistique* 467 (1): 61–89.

Barnett, W. S. 2011. “Effectiveness of Early Educational Intervention.” *Science* 333 (6045): 975–78.

Baronnet, Juliette, Alice Best, Florence Brunet, and Nicolas Duvoux. 2021. “Accompagner Les Familles Monoparentales Moyens Et Enjeux de l’autonomisation de Publics à La Croisée Des Vulnérabilités.” Dossier d'étude 225. Paris: Cnaf.

Barr, Andrew, Jonathan Eggleston, and Alexander A Smith. 2022. “Investing in Infants: The Lasting Effects of Cash Transfers to New Families*.” *The Quarterly Journal of Economics*, April, qjac023.

Bauernschuster, Stefan, and Martin Schlotter. 2015. “Public Child Care and Mothers’ Labor Supply—Evidence from Two Quasi-Experiments.” *Journal of Public Economics* 123 (March): 1–16.

Bazen, Stephen, Joutard Xavier, and Hélène Périvier. 2022. “Measuring the Child Penalty Early in a Career: The Case of Young Adults in France.” *SSRN Electronic Journal*.

Bedard, Kelly, and Olivier Deschênes. 2005. “Sex Preferences, Marital Dissolution, and the Economic Status of Women.” *Journal of Human Resources* XL (2): 411–34.

Behaghel, Luc, Bruno Crépon, and Marc Gurgand. 2014. “Private and Public Provision of Counseling to Job Seekers: Evidence from a Large Controlled Experiment.” *American Economic Journal: Applied Economics* 6 (4): 142–74.

Bergemann, Annette, and Gerard J. Van Den Berg. 2008. “Active Labor Market Policy Effects for Women in Europe — A Survey.” *Annales d’Économie Et de Statistique*, no. 91/92: 385–408.

Berger, Lawrence M., Lidia Panico, and Anne Solaz. 2021. “The Impact of Center-Based Childcare Attendance on Early Child Development: Evidence From the French Elfe Cohort.” *Demography*, February, 8977274.

Bergman, Peter, Raj Chetty, Stefanie DeLuca, Nathaniel Hendren, Lawrence Katz, and Christopher Palmer. 2019. “Creating Moves to Opportunity: Experimental Evidence on Barriers to Neighborhood Choice.” w26164. Cam-

bridge, MA: National Bureau of Economic Research.
Bernal, Raquel, and Michael P. Keane. 2010. “Quasi-Structural Estimation of a Model of Childcare Choices and Child Cognitive Ability Production.” *Journal of Econometrics* 156 (1): 164–89.
Bertanha, Marinho, Carolina Caetano, Hugo Jales, and Nathan Seegert. 2023. “Bunching Estimation Methods.” In *Handbook of Labor, Human Resources and Population Economics*.
Bertrand, Marianne. 2011. “New Perspectives on Gender.” In *Handbook of Labor Economics*, 4:1543–90. Elsevier.
Bertrand, Marianne, Patricia Cortes, Claudia Olivetti, and Jessica Pan. 2021. “Social Norms, Labour Market Opportunities, and the Marriage Gap Between Skilled and Unskilled Women.” *The Review of Economic Studies* 88 (4): 1936–78.
Bertrand, Marianne, Claudia Goldin, and Lawrence F Katz. 2010. “Dynamics of the Gender Gap for Young Professionals in the Financial and Corporate Sectors.” *American Economic Journal: Applied Economics* 2 (3): 228–55.
Bertrand, Marianne, Emir Kamenica, and Jessica Pan. 2015. “Gender Identity and Relative Income Within Households*.” *The Quarterly Journal of Economics* 130 (2): 571–614.
Bessière, Céline, and Sibylle Gollac. 2020. “1. La Famille, Une Institution Économique.” In *Le Genre Du Capital*, 23–51. SH / l’envers Des Faits. Paris: La Découverte.
Betkó, János. 2023. “Effects of Welfare Policies Based on Autonomy and Unconditionality.” PhD thesis, Radboud University.
Bhargava, Saurabh, and Dayanand Manoli. 2015. “Psychological Frictions and the Incomplete Take-Up of Social Benefits: Evidence from an IRS Field Experiment.” *American Economic Review* 105 (11): 3489–529.
Bianchi, Suzanne, Laurent Lesnard, Tiziana Nazio, and Sara Raley. 2014. “Gender and Time Allocation of Cohabiting and Married Women and Men in France, Italy, and the United States.” *Demographic Research* 31 (July): 183–216.
Bierbrauer, Felix J., Pierre C. Boyer, and Andreas Peichl. 2021. “Politically Feasible Reforms of Nonlinear Tax Systems.” *The American Economic Review* 111 (1): 153–91.
Binder, Ariel J., and John Bound. 2019. “The Declining Labor Market Prospects of Less-Educated Men.” *Journal of Economic Perspectives* 33 (2): 163–90.
Binns, Reuben, Max Van Kleek, Michael Veale, Ulrik Lyngs, Jun Zhao, and Nigel Shadbolt. 2018. “’It’s Reducing a Human Being to a Percentage’; Perceptions of Justice in Algorithmic Decisions.” *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI ’18*, 1–14.
Bitler, Marianne, and Madeline Zavodny. 2002. “Child Abuse and Abortion Availability.” *American Economic Review* 92 (2): 363–67.
Bjerre-Nielsen, Andreas, Lykke Sterll Christensen, Mikkel Høst Gandil, and Hans Henrik Sievertsen. 2023. “Playing the System: Address Manipulation and Access to Schools.” May 30, 2023.
Blanchflower, David G., and Andrew E. Clark. 2021. “Children, Unhappiness and Family Finances.” *Journal of Population Economics* 34 (2): 625–53.
Blanden, Jo, Emilia Del Bono, Kirstine Hansen, and Birgitta Rabe. 2022. “Quantity and Quality of Childcare and Children’s Educational Outcomes.” *Journal of Population Economics* 35 (2): 785–828.
Blandhol, Christine, Magne Mogstad, John Bonney, and Alexander Torgovitsky. 2022. “When Is TSLS Actually LATE?” Becker Friedman Institute for Economics at UChucago. February 9, 2022.
Blau, Francine D., and Lawrence M. Kahn. 2017. “The Gender Wage Gap: Extent, Trends, and Explanations.” *Journal of Economic Literature* 55 (3): 789–865.
Blomquist, Soren, and Whitney K. Newey. 2017. “The Bunching Estimator Cannot Identify the Taxable Income Elasticity.” SSRN Scholarly Paper. Rochester, NY. November 21, 2017.
Bloom, Howard S., Carolyn J. Hill, and James A. Riccio. 2003. “Linking Program Implementation and Effectiveness: Lessons from a Pooled Sample of Welfare-to-Work Experiments.” *Journal of Policy Analysis and Management* 22 (4): 551–75.
Blundell, Richard, Mike Brewer, Peter Haan, and Andrew Shephard. 2009. “Optimal Income Taxation of Lone Mothers: An Empirical Comparison of the UK and Germany.” *The Economic Journal* 119 (535): F101–21.
Blundell, Richard, Monica Costa Dias, Costas Meghir, and Jonathan Shaw. 2016. “Female Labor Supply, Human Capital, and Welfare Reform.” *Econometrica* 84 (5): 1705–53.
Bodson, Lucile, and Mathias Kuépié. 2012. “Effet de La Dynamique Des Revenus Professionnels Des Conjoints Sur La Dynamique Matrimoniale Au Luxembourg.” *Revue Économique* 63 (2): 261–81.
Bohren, J. Aislinn, Peter Hull, and Alex Imas. 2022. “Systemic Discrimination: Theory and Measurement.” Work-

ing Paper. Working Paper Series. March 2022.
Boneva, Teodora, Marta Golin, Katja Kaufmann, and Christopher Rauh. 2022. “Beliefs about Maternal Labor Supply.” *IZA Discussion Papers*.
Bonnet, Carole, Bertrand Garbinti, and Anne Solaz. 2021. “The Flip Side of Marital Specialization: The Gendered Effect of Divorce on Living Standards and Labor Supply.” *Journal of Population Economics* 34 (2): 515–73.
———. 2022. “Does Part-Time Mothering Help Get a Job? The Role of Shared Custody in Women’s Employment.” *European Journal of Population* 38 (December): 885–913.
Bonnet, Carole, Marine Montaignac, and Anne Solaz. 2024. “Séparation Des Parents : Quel Impact Sur Le Niveau de Vie Des Enfants ?” Note d'analyse 1. Paris: France Stratégie.
Bono, Emilia Del, Marco Francesconi, Yvonne Kelly, and Amanda Sacker. 2016. “Early Maternal Time Investment and Early Child Outcomes.” *The Economic Journal* 126 (596): F96–135.
Bono, Pierre-Henri, Rozenn Desplatz, Sandra Debu, and Clément Lacouette-Fougère. 2021. “Le Lent Développement Des Évaluations d’impact En France : Une Approche Par Les Acteurs.” *Revue Francaise d’administration Publique* N° 177 (1): 17–28.
Bonoli, G. 2012. “Active Labour Market Policy and Social Investment: A Changing Relationship.” In *Towards a Social Investment Welfare State ? Ideas, Policies and Challenges*, Morel, N., Palier, B. & Palme, J., 181–204. Bristol: Policy Press.
Bonoli, Giuliano. 2010. “The Political Economy of Active Labor-Market Policy.” *Politics & Society* 38 (4): 435–57.
———. 2011. “Active Labour Market Policy and Social Investment: A Changing Relationship,” November.
Bonoli, Giuliano, and Frank Reber. 2010. “The Political Economy of Childcare in OECD Countries: Explaining Cross-National Variation in Spending and Coverage Rates.” *European Journal of Political Research* 49 (1): 97–118.
Borusyak, Kirill, UC Berkeley, and Peter Hull. 2023. “Design-Based Identification with Formula Instruments: A Review.”
Borusyak, Kirill, Peter Hull, and Xavier Jaravel. 2022. “Quasi-Experimental Shift-Share Research Designs.” Edited by Dirk Krueger. *The Review of Economic Studies* 89 (1): 181–213.
Borusyak, Kirill, Xavier Jaravel, and Jann Spiess. 2022. “Revisiting Event Study Designs: Robust and Efficient Estimation.” April 19, 2022.
Bose, Gautam, Tarun Jain, and Sarah Walker. 2022. “Women’s Labor Force Participation and Household Technology Adoption.” *European Economic Review* 147 (August): 104181.
Bourguignon, François. 2009. “Rapport Final Sur l’évaluation Des Expérimentations rSa - Comité d’evaluation Des Expérimentations.” Paris: Haut Commissaire aux Solidarités actives contre la Pauvreté.
Bousselin, Audrey. 2022. “Access to Universal Childcare and Its Effect on Maternal Employment.” *Review of Economics of the Household* 20 (2): 497–532.
Bouteillec, Nathalie Le, Lamia Kandil, and Anne Solaz. 2014b. “L’accueil En Crèche En France : Quels Enfants y Ont Accès ?” *Population Societes* N° 514 (8): 1–4.
———. 2014a. “L’accueil En Crèche En France : Quels Enfants y Ont Accès ?” *Population Societes* 514 (8): 1–4.
Bozio, Antoine. 2014. “L’évaluation Des Politiques Publiques : Enjeux, Méthodes Et Institutions.” *Revue Francaise d’economie* Volume XXIX (4): 59–85.
Bozio, Antoine, Thomas Breda, and Malka Guillot. 2023. “Using Payroll Taxes as a Redistribution Tool.” *Journal of Public Economics* 226 (October): 104986.
Bozio, Antoine, Brice Fabre, Claire Leroy, Nolwenn Loisel, and Maxime Tô. 2023. “La Réforme de 2019 de La Prime d’activité : Analyse Sur La Possibilité de Mesurer Les Effets Sur l’emploi.” Paris: France stratégie.
Brady, Michelle, and Kay Cook. 2015. “The Impact of Welfare to Work on Parents and Their Children.” *Evidence Base* 3: 23.
Breton, Didier, Nicolas Belliot, Magali Barbieri, Hippolyte d’Albis, and Magali Mazuy. 2021. “L’évolution Démographique Récente de La France. Moins de Naissances, de Mariages Et de Migrations, Plus de Décès… La Covid-19 Bouleverse La Dynamique de La Population Française.” *Population* 76 (4): 577–638.
Briard, Karine. 2020. “L’élasticité de l’offre de Travail Des Femmes En France. Petite Revue de Méthodes Et de Résultats.” *Revue de l’OFCE* 169 (5): 39–72.
Briselli, Giulia, and Libertad Gonzalez. 2023. “Are Men’s Attitudes Holding Back Fertility and Women’s Careers? Evidence from Europe.” NBER Working Paper. May 19, 2023.
Brodeur, Abel, Nikolai Cook, and Anthony Heyes. 2020. “Methods Matter: P-Hacking and Publication Bias in

Causal Analysis in Economics." *American Economic Review* 110 (11): 3634–60.
Broussard, C. Anne. 2010. "Research Regarding Low-Income Single Mothers' Mental and Physical Health: A Decade in Review." *Journal of Poverty* 14 (4): 443–51.
Brown, Benjamin P., Luciana E. Hebert, Melissa Gilliam, and Robert Kaestner. 2020. "Distance to an Abortion Provider and Its Association with the Abortion Rate: A Multistate Longitudinal Analysis." *Perspectives on Sexual and Reproductive Health* 52 (4): 227–34.
Brown, Elizabeth. 2020. "Parcours Conjugaux, Violence Conjugale Et Différences de Genre." In *Violences Et Rapports de Genre*, by Elizabeth Brown, Alice Debauche, Christelle Hamel, and Magali Mazuy, 38. Grandes Enquêtes.
Brück, Tilman, John P. Haisken-De New, and Klaus F. Zimmermann. 2006. "Creating Low Skilled Jobs by Subsidizing Market-Contracted Household Work." *Applied Economics* 38 (8): 899–911.
Büchel, Felix, and Antje Mertens. 2004. "Overeducation, Undereducation, and the Theory of Career Mobility." *Applied Economics* 36 (8): 803–16.
Bullock, Heather E., Wendy R. Williams, and Wendy M. Limbert. 2003. "Predicting Support for Welfare Policies: The Impact of Attributions and Beliefs About Inequality." *Journal of Poverty* 7 (3): 35–56.
Burger, Kaspar. 2010. "How Does Early Childhood Care and Education Affect Cognitive Development? An International Review of the Effects of Early Interventions for Children from Different Social Backgrounds." *Early Childhood Research Quarterly* 25 (2): 140–65.
Burtless, Gary, and Jerry A. Hausman. 1978. "The Effect of Taxation on Labor Supply: Evaluating the Gary Negative Income Tax Experiment." *Journal of Political Economy* 86 (6): 1103–30.
Caenen, Yann, and Pauline Virot. 2023. "La Part Des Enfants de Moins de 3 Ans Confiés Principalement à Une Assistante Maternelle Ou Une Crèche a Presque Doublé Entre 2002 Et 2021." *DREES*, Études et Résultats, no. 1257 (February).
Cahuc, Pierre. 2002. "A Quoi Sert La Prime Pour l'emploi ?" *Revue Française d'économie* 16 (3): 3–61.
Cahuc, Pierre, Stéphane Carcillo, and André Zylberberg. 2014. *Labor Economics*. Second Edition. Cambridge, MA: MIT Press.
Cairo, Sofie, and Robert Mahlstedt. 2023. "The Disparate Effects of Information Provision: A Field Experiment on the Work Incentives of Social Welfare." *Journal of Public Economics* 226 (October): 104987.
Caldwell, Sydnee, Scott Nelson, and Daniel Waldinger. 2023. "Tax Refund Uncertainty: Evidence and Welfare Implications." *American Economic Journal: Applied Economics* 15 (2): 352–76.
Callaway, Brantly, and Pedro H. C. Sant'Anna. 2020a. "Difference-in-Differences with Multiple Time Periods." *Journal of Econometrics*, December, S0304407620303948.
———. 2020b. "Difference-in-Differences with Multiple Time Periods." *Journal of Econometrics*, December, S0304407620303948.
Calonico, S., M. D. Cattaneo, and R. Titiunik. 2014. "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs." *Econometrica* 82 (6): 2295–2326.
Camilli, Gregory, Sadako Vargas, Sharon Ryan, and W Steven Barnett. 2010. "Meta-Analysis of the Effects of Early Education Interventions on Cognitive and Social Development." *Teachers College Record* 112 (3): 579–620.
Campbell, Mhairi, Hilary Thomson, Candida Fenton, and Marcia Gibson. 2016b. "Lone Parents, Health, Wellbeing and Welfare to Work: A Systematic Review of Qualitative Studies." *BMC Public Health* 16 (1): 188.
———. 2016a. "Lone Parents, Health, Wellbeing and Welfare to Work: A Systematic Review of Qualitative Studies." *BMC Public Health* 16 (1): 188.
Cantillon, Bea. 2011. "The Paradox of the Social Investment State: Growth, Employment and Poverty in the Lisbon Era." *Journal of European Social Policy* 21 (5): 432–49.
Cantillon, Bea, and Wim Van Lancker. 2011. "Le Paradoxe de l'état d'investissement Social : Pourquoi La Pauvreté n'a-t-Elle Pas Baissé ?" *Reflets Et Perspectives de La Vie Économique* L (4): 55–73.
Carayon, Lisa. 2018. "Enfants Délaissés : Mères Sans-Papiers. Pour Une Critique Nécessaire de l'article 30 Du Projet de Loi Sur l'asile Et l'immigration (Second Épisode)." *La Revue Des Droits de l'homme. Revue Du Centre de Recherches Et d'études Sur Les Droits Fondamentaux*, April.
Carbuccia, Laudine, Carlo Barone, Gregoire Borst, Angela Greulich, Lidia Panico, and Maxime Tô. 2020. "Revue de Littérature Sur Les Politiques d'accompagnement Au Développement Des Capacités Des Jeunes Enfants." Dossier d'étude 215. CNAF.
Carbuccia, Laudine, Carlo Barone, Coralie Chevallier, and Arthur Heim. 2023. "Can Information and Assistance Increase Application and Access to Formal Childcare: Experimental Evidence from France." Pre-registration

analysis. October 22, 2023.
Carbuccia, Laudine, Valentin Thouzeau, Carlo Barone, and Coralie Chevallier. 2023. "Unequal Access to Early Childcare: What Role Do Demand-Side Factors Play? A PRISMA Systematic Review." LIEPP Working Paper. Paris. January 2023.
Carcasson, Martín. 2006. "Ending Welfare as We Know It: President Clinton and the Rhetorical Transformation of the Anti-Welfare Culture." *Rhetoric and Public Affairs* 9 (4): 655–92.
Card, David, and Dean R. Hyslop. 2005. "Estimating the Effects of a Time-Limited Earnings Subsidy for Welfare-Leavers." *Econometrica* 73 (6): 1723–70.
Card, David, Jochen Kluve, and Andrea Weber. 2010. "Active Labour Market Policy Evaluations: A Meta□Analysis." *The Economic Journal* 120 (548): F452–77.
———. 2018. "What Works? A Meta Analysis of Recent Active Labor Market Program Evaluations." *Journal of the European Economic Association* 16 (3): 894–931.
Carlsson, Sissa, and Stephan L Thomsen. 2015. "Improving the Allocation of Spots in Child Care Facilities for Toddlers in Germany: A Mechanism Design Approach." IZA Discussion papers 8976. IZA.
Cartier, Marie, Anaïs Collet, Estelle Czerny, Pierre Gilbert, Marie-Hélène Lechien, and Sylvie Monchatre. 2017. "Pourquoi Les Parents Préfèrent-Ils La Crèche ? Les Représentations Hierarchisées Des Modes de Garde Professionnels." *Revue Française Des Affaires Sociales*, no. 2: 250–65.
Case, Anne, Angela Fertig, and Christina Paxson. 2005. "The Lasting Impact of Childhood Health and Circumstance." *Journal of Health Economics* 24 (2): 365–89.
Cassar, Lea, and Stephan Meier. 2018. "Nonmonetary Incentives and the Implications of Work as a Source of Meaning." *Journal of Economic Perspectives* 32 (3): 215–38.
Castell, Laura, Marc Gurgand, Clement Imbert, and Todor Tochev. 2022. "Take-up of Social Benefits: Experimental Evidence from France." December 2022.
Cattaneo, Matias D. 2010. "Efficient Semiparametric Estimation of Multi-Valued Treatment Effects Under Ignorability." *Journal of Econometrics* 155 (2): 138–54.
Cattaneo, Matias D., Michael Jansson, and Xinwei Ma. 2021. "Local Regression Distribution Estimators." *Journal of Econometrics*, March, 105074.
Cattaneo, Matias D, and Rocıo Titiunik. 2022. "Regression Discontinuity Designs." *Annual Review of Economics*, 50.
Cattoen, Victoire, Clotilde Leire, Nathanaël Trellet, and Izaya Youssouf. 2022. "Parcours de Vie Et Expériences Du RSA. Une Approche Sociologique." Université de Lille.
Cavaillé, Charlotte, and Karine Van der Straeten. 2023. "Immigration and Support for Redistribution: Lessons from Europe." *Journal of Economic Literature* 61 (3): 958–76.
Cavapozzi, Danilo, Marco Francesconi, and Cheti Nicoletti. 2021. "The Impact of Gender Role Norms on Mothers' Labor Supply." *Journal of Economic Behavior and Organization* 186 (June): 113–34.
Celhay, Pablo A., Bruce D. Meyer, and Nikolas Mittag. 2022. "Stigma in Welfare Programs." Working Paper. Working Paper Series. July 2022.
Cervera, Melaine, Céline Émond, Renaud Hourcade, Céline Jung, and Rémi Le Gall. 2017. "L'accompagnement Social Et Ses Effets Sur Les Bénéficiaires." Revue de littérature critique 195. Caisse nationale des allocations familiales.
Chachou, Khadija. 2019. "Le Non-Recours Au Droit à l'accompagnement Des Publics Allocataires Du RSA: Analyse Des Facteurs de La Non-Participation." Mémoire de Master. Nancy: Université de Lorraine.
Chaisemartin, C De, and Xavier D'Haultfoeuille. 2017. "Fuzzy Di□erences-in-Di□erences." *Review of Economic Studies*, no. 1: 1–30.
Chaisemartin, Clément De, and Xavier D'Haultfœuille. 2017. "Fuzzy Differences-in-Differences." *The Review of Economic Studies*, 30.
Chaisemartin, Clément de, Quentin Daviot, Marc Gurgand, and Sophie Kern. 2021. "Lutter Contre Les Inégalités Dès La Petite Enfance: Évaluation à Grande Échelle Du Programme Parler Bambin." PSE Working Paper. HAL.
Chaisemartin, Clément de, and Jaime Ramirez-Cuellar. 2022. "At What Level Should One Cluster Standard Errors in Paired and Small-Strata Experiments?" w27609. Cambridge, MA: National Bureau of Economic Research.
Champagne, Clara, Ariane Pailhé, and Anne Solaz. 2015b. "Le Temps Domestique Et Parental Des Hommes Et Des Femmes : Quels Facteurs d'évolutions En 25 Ans ?" *Economie Et Statistique* 478 (1): 209–42.
———. 2015a. "Le Temps Domestique Et Parental Des Hommes Et Des Femmes : Quels Facteurs d'évolutions

En 25 Ans ?” *Economie Et Statistique* 478 (1): 209–42.
Chan, Marc K, and Robert Moffitt. 2018. “Welfare Reform and the Labor Market.” *Annual Review of Economics* 10: 347–81.
Chancel, Lucas, and Thomas Piketty. 2021. “Global Income Inequality, 1820–2020: The Persistence and Mutation of Extreme Inequality.” *Journal of the European Economic Association* 19 (6): 3025–62.
Chappell, Marisa. 2010. *The War on Welfare: Family, Poverty, and Politics in Modern America*. University of Pennsylvania Press.
Chareyron, Sylvain. 2018. “Pauvreté Et Non-Recours Aux Dispositifs Sociaux : L’étude Du RSA « Socle Seul ».” *Économie & Prévision* 213 (1): 41–59.
Chareyron, Sylvain, David Gray, and Yannick L’Horty. 2018. “Raising Take-up of Social Assistance Benefits Through a Simple Mailing: Evidence from a French Field Experiment.” *Revue d’économie Politique* 128 (5): 777–805.
Chareyron, Sylvain, Rémi Le Gall, and Yannick L’Horty. 2022. “Droits Et Devoirs Du RSA : L’impact Des Contrôles Sur La Participation Des Bénéficiaires:” *Revue Économique* Vol. 73 (5): 735–59.
Chen, Yan, and YingHua He. 2021. “Information Acquisition and Provision in School Choice: An Experimental Study.” *Journal of Economic Theory* 197 (October): 105345.
Cheng, Chien-Tzu. 2022. “The Impact of Home Pregnancy Testing on Fertility and Women’s Later-Life Outcomes.” Brown University. December 2022.
Chernozhukov, Victor, and Christian Hansen. 2008. “Instrumental Variable Quantile Regression: A Robust Inference Approach.” *Journal of Econometrics* 142 (1): 379–98.
Chetty, Raj, John N. Friedman, and Emmanuel Saez. 2013. “Using Differences in Knowledge Across Neighborhoods to Uncover the Impacts of the EITC on Earnings.” *The American Economic Review* 103 (7): 2683–2721.
Chetty, Raj, and Emmanuel Saez. 2013. “Teaching the Tax Code: Earnings Responses to an Experiment with EITC Recipients.” *American Economic Journal: Applied Economics* 5 (1): 1–31.
Chetty, R., J. N. Friedman, T. Olsen, and L. Pistaferri. 2011. “Adjustment Costs, Firm Responses, and Micro Vs. Macro Labor Supply Elasticities: Evidence from Danish Tax Records.” *The Quarterly Journal of Economics* 126 (2): 749–804.
Chiappori, Pierre-André. 2017. *Matching with Transfers: The Economics of Love and Marriage*. The Gorman Lectures in Economics. Princeton: Princeton University Press.
Chiappori, Pierre-Andre, and Maurizio Mazzocco. 2017. “Static and Intertemporal Household Decisions.” *Journal of Economic Literature* 55 (3): 985–1045.
Chiappori, Pierre-André, Natalia Radchenko, and Bernard Salanié. 2018. “Divorce and the Duality of Marital Payoff.” *Review of Economics of the Household* 16: 833–58.
Chung, Heejung. 2020. “Gender, Flexibility Stigma and the Perceived Negative Consequences of Flexible Working in the UK.” *Social Indicators Research* 151 (2): 521–45.
Clark, Andrew E., Conchita D’Ambrosio, and Anthony Lepinteur. 2022. “Marriage as Insurance: Job Protection and Job Insecurity in France.” *Review of Economics of the Household*, December.
Cnesco. 2016. “Inégalités Sociales Et Migratoire: Comment l’école Amplifie-t’elle Les Inégalités ?” Rapport. Paris: Conseil national d’évaluation du système scolaire.
Coile, Courtney C., and Mark G. Duggan. 2019. “When Labor’s Lost: Health, Family Life, Incarceration, and Education in a Time of Declining Economic Opportunity for Low-Skilled Men.” *Journal of Economic Perspectives* 33 (2): 191–210.
Coles, Jeffrey L., Elena Pa□, Nathan Seegert, and Matthew Smith. 2022. “How Do Firms Respond to Corporate Taxes?” *Journal of Accounting Research* 60 (3): 965–1006.
Colquitt, Jason A., Donald E. Conlon, Michael J. Wesson, Christopher O. L. H. Porter, and K. Yee Ng. 2001. “Justice at the Millennium: A Meta-Analytic Review of 25 Years of Organizational Justice Research.” *Journal of Applied Psychology* 86 (3): 425–45.
Combe, Julien, and Arthur Heim. 2024. “Rage Against the Matching: Fairness and Inequalities in a Market Design Experiment of Daycare Assignments in France.” In *Social Investment & The Changing Face of Poverty: Essays on the Design and Evaluation of Family and Social Policies in France*, Paris School of Economics. 1. Paris, France: PhD Thesis.
Combe, Julien, and Jan Christoph Schlegel. 2021. “Reallocation with Priorities.” SSRN Scholarly Paper. Rochester, NY. May 6, 2021.
Combe, Julien, Olivier Tercieux, and Camille Terrier. 2022. “The Design of Teacher Assignment: Theory and

Evidence." *The Review of Economic Studies* 89 (6): 3154–3222.
Commission, European. 2015. "Social Investment - Employment, Social Affairs & Inclusion -." https://ec.europa.eu/social. 2015.
Connelly, Rachel, and Jean Kimmel. 2003. "The Effect of Child Care Costs on the Employment and Welfare Recipiency of Single Mothers." *Southern Economic Journal* 69 (3): 498–519.
Cook, Kay E. 2012. "Social Support in Single Parents' Transition from Welfare to Work: Analysis of Qualitative Findings1." *International Journal of Social Welfare* 21 (4): 338–50.
Corak, Miles. 2013. "Income Inequality, Equality of Opportunity, and Intergenerational Mobility." *Journal of Economic Perspectives* 27 (3): 79–102.
Cortés, Patricia, and José Tessada. 2011. "Low-Skilled Immigration and the Labor Supply of Highly Skilled Women." *American Economic Journal: Applied Economics* 3 (3): 88–123.
Costemalle, Vianney. 2017. "How Long Do Situations of Single Parenthood Last? An Estimation Based on French Data." *Economie Et Statistique / Economics and Statistics*, no. 493 (July): 87–112.
Cour des comptes. 2022. "Le Revenu de Solidarité Active." Paris: Cours des comptes.
Craig, Lyn, and Killian Mullan. 2011. "How Mothers and Fathers Share Childcare: A Cross-National Time-Use Comparison." *American Sociological Review* 76 (6): 834–61.
Crepon, Bruno, Marc Gurgand, Thierry Kamionka, and Laurent Lequien. 2013. "Is Counseling Welfare Recipients Cost-E□ective? Lessons from a Random Experiment." Série Des Documents de Travail. CREST. February 2013.
Crépon, Bruno, and Gerard J. van den Berg. 2016b. "Active Labor Market Policies." *Annual Review of Economics* 8 (1): 521–46.
———. 2016a. "Active Labor Market Policies." *Annual Review of Economics* 8 (1): 521–46.
Cunha, F., and J. J. Heckman. 2007. "The Technology of Skill Formation." *American Economic Review* 97 (2): 31–47.
Cunha, Flavio, James J. Heckman, Lance Lochner, and Dimitriy V. Masterov. 2006. "Interpreting the Evidence on Life Cycle Skill Formation." In *Handbook of the Economics of Education*, 1:697–812. Elsevier.
Currie, Janet, and Douglas Almond. 2011. "Human Capital Development Before Age Five." In *Handbook of Labor Economics*, 4:1315–1486. Elsevier.
Dahl, Gordon B., Runjing Lu, and William Mullins. 2022. "Partisan Fertility and Presidential Elections." *American Economic Review: Insights* 4 (4): 473–90.
Dahl, Gordon B., and Enrico Moretti. 2008. "The Demand for Sons." *Review of Economic Studies* 75 (4): 1085–1120.
Damon, Julien. 2015. "L'investissement Social : Contenu Et Portée d'une Notion En Vogue." *Revue de Droit Sanitaire Et Social*, no. 4: 722–33.
———. 2018. "La Juste Prestation : Pour Des Prestations Et Un Accompagnement Ajustés." Rapport au 1er Ministre.
Dardier, Agathe, Quynh-Chi Doan, and Cléo Lhermet. 2022. "La Revalorisation Du Bonus Individuel En 2019 a Fortement Élargi Le Champ Des Bénéficiaires de La Prime d'activité." 1225. Paris: DREES.
DARES. 2021. "Expérimentation Territoire Zéro Chômeurs de Longue Durée." Paris: DARES.
Darity, William A., and Patrick L. Mason. 1998. "Evidence on Discrimination in Employment: Codes of Color, Codes of Gender." *Journal of Economic Perspectives* 12 (2): 63–90.
De Deken, Johan. 2016. "Conceptualiser Et Mesurer l'investissement Social Dans Une Perspective Comparative." *Informations Sociales* n° 192 (1): 28.
Deeming, Christopher, and Paul Smyth. 2015. "Social Investment After Neoliberalism: Policy Paradigms and Political Platforms." *Journal of Social Policy* 44 (2): 297–318.
Défenseur des droits. 2017. "Lutte Contre La Fraude Aux Prestations Sociales : À Quel Prix Pour Les Droits Des Usagers ?" Paris: Défenseur des droits.
Del Boca, Daniela, Christopher Flinn, and Matthew Wiswall. 2014. "Household Choices and Child Development." *The Review of Economic Studies* 81: 137–85.
Delacrétaz, David, Scott Duke Kominers, and Alexander Teytelboym. 2023. "Matching Mechanisms for Refugee Resettlement." *American Economic Review*, no. 10 (October): 2689–2717.
Delahais, Thomas, and Clément Lacouette-Fougère. 2019. "Try Again. Fail Again. Fail Better. Analysis of the Contribution of 65 Evaluations to the Modernisation of Public Action in France." *Evaluation* 25 (2): 131–48.
Delattre, Eric, Richard K. Moussa, and Mareva Sabatier. 2019. "Health Condition and Job Status Interactions:

Econometric Evidence of Causality from a French Longitudinal Survey.” *Health Economics Review* 9 (1): 3.
Désesquelles, Aline, and Annie Kensey. 2006. “Les Detenus Et Leur Famille : Des Liens Presque Toujours Maintenus Mais Parfois Tres Distendus.” *Données Sociales*, Population, famille, 59–67.
Desplatz, Rozenn, and Clément Lacouette Fougère. 2019. “L’évaluation Des Politiques Publiques En France.” Document de travail 2019-13. Paris: France Stratégie.
Dherbecourt, Clément, and Jean Flamand. 2023. “Inégalité Des Chances : Ce Qui Compte Le Plus.” Note d'analyse 120. Paris: France Stratégie.
Diamond, Peter A. 1998. “Optimal Income Taxation: An Example with a U-Shaped Pattern of Optimal Marginal Tax Rates.” *The American Economic Review* 88 (1): 83–95.
Ding, Peng, Avi Feller, and Luke Miratrix. 2019. “Decomposing Treatment Effect Variation.” *Journal of the American Statistical Association* 114 (525): 304–17.
Domingo, Pauline, and Muriel Pucci. 2013. “Les Vecteurs Du Non-Recours Au Revenu de Solidarité Active Du Point de Vue de l’usager.” *Informations Sociales* 178 (4): 72–80.
———. 2014. “Impact Du Non-Recours Sur l’efficacité Du RSA Activité Seul.” *Economie Et Statistique* 467 (1): 117–40.
Donohue, John J., III, and Steven D. Levitt. 2008. “Measurement Error, Legalized Abortion, and the Decline in Crime: A Response to Foote and Goetz*.” *The Quarterly Journal of Economics* 123 (1): 425–40.
Donohue, John J., and Steven D Levitt. 2001. “THE IMPACT OF LEGALIZED ABORTION ON CRIME.” *QUARTERLY JOURNAL OF ECONOMICS* 141 (2).
Donohue, John J., and Steven D. Levitt. 2004. “Further Evidence That Legalized Abortion Lowered Crime: A Reply to Joyce.” *The Journal of Human Resources* 39 (1): 29–49.
———. 2019. “The Impact of Legalized Abortion on Crime over the Last Two Decades.” Working Paper. Working Paper Series. May 2019.
Dotti Sani, Giulia M., and Judith Treas. 2016. “Educational Gradients in Parents’ Child-Care Time Across Countries, 1965-2012: Educational Gradients in Parents’ Child-Care Time.” *Journal of Marriage and Family* 78 (4): 1083–96.
Doyle, Orla. 2020. “The First 2,000 Days and Child Skills.” *Journal of Political Economy* 128 (6): 2067–2122.
Drange, Nina, and Tarjei Havnes. 2015. “Child Care Before Age Two and the Development of Language and Numeracy. Evidence from a Lottery.” Discussion paper 808. Statistics Norway, Research Department.
DREES. 2021. “Minima Sociaux Et Prestations Sociales Ménages Aux Revenus Modestes Et Redistribution.” PANORAMAS DE LA DREES SOCIAL.
———. 2022. “Minima Sociaux Et Prestations Sociales Ménages Aux Revenus Modestes Et Redistribution.” PANORAMAS DE LA DREES SOCIAL.
Dubins, D E, and D. A. Freedman. 1981. “MACHIAVELLI AND THE GALE-SHAPLEY ALGORITHM.” *The American Mathematical Monthly* 88 (7): 485–94.
Dubois, Vincent. 2021. *Contrôler Les Assistés. Genèses Et Usages d’un Mot d’ordre*. Raisons d’agir.
Duflo, Esther. 2017. “The Economist as Plumber.” *American Economic Review* 107 (5): 1–26.
Duncan, Greg J., Katherine Magnuson, Ariel Kalil, and Kathleen Ziol-Guest. 2012a. “The Importance of Early Childhood Poverty.” *Social Indicators Research* 108 (1): 87–98.
———. 2012b. “The Importance of Early Childhood Poverty.” *Social Indicators Research* 108 (1): 87–98.
Duncan, Greg J., Kathleen M. Ziol-Guest, and Ariel Kalil. 2010. “Early-Childhood Poverty and Adult Attainment, Behavior, and Health.” *Child Development* 81 (1): 306–25.
Duncan, Greg, Ariel Kalil, Magne Mogstad, and Mari Rege. 2023. “Chapter 1 - Investing in Early Childhood Development in Preschool and at Home.” In *Handbook of the Economics of Education*, edited by Eric A. Hanushek, Stephen Machin, and Ludger Woessmann, 6:1–91. Elsevier.
Duncan, Otis Dudley, and Beverly Duncan. 1955. “A Methodological Analysis of Segregation Indexes.” *American Sociological Review* 20 (2): 210.
Duncan, Simon, and Rosalind Edwards. 1997. “Lone Mothers and Paid Work - Rational Economic Man or Gendered Moral Rationalities?” *Feminist Economics* 3 (2): 29–61.
Dur, Umut, Scott Duke Kominers, Parag A. Pathak, and Tayfun Sönmez. 2018. “Reserve Design: Unintended Consequences and the Demise of Boston’s Walk Zones.” *Journal of Political Economy* 126 (6): 2457–79.
Duriancik, David M, and Courtney R Goff. 2019. “Children of Single-Parent Households Are at a Higher Risk of Obesity: A Systematic Review.” *Journal of Child Health Care* 23 (3): 358–69.
Durlauf, Steven N. 2008. “Affirmative Action, Meritocracy, and Efficiency.” *Politics, Philosophy & Economics* 7

(2): 131–58.
Durlauf, Steven N., Andros Kourtellos, and Chih Ming Tan. 2022. “The Great Gatsby Curve.” Becker Friedman Institute for Economics. Chicago. February 2022.
Durovic, Anja. 2023. “Rising Electoral Fragmentation and Abstention: The French Elections of 2022.” *West European Politics* 46 (3): 614–29.
Duvoux, Nicolas, and Michèle Lelièvre. 2021. “La Pauvreté Démultipliée Dimensions, Processus Et Réponses Printemps 2020 | Printemps 2021.” Rapport au premier ministre. Conseil national des politiques de lutte contre la pauvreté et l’exclusion sociale.
Dwyer, Peter, Lisa Scullion, Katy Jones, Jenny McNeill, and Alasdair B. R. Stewart. 2020. “Work, Welfare, and Wellbeing: The Impacts of Welfare Conditionality on People with Mental Health Impairments in the UK.” *Social Policy & Administration* 54 (2): 311–26.
Ebert, Udo. 2005. “Optimal Anti□Poverty Programmes: Horizontal Equity and the Paradox of Targeting.” *Economica* 72 (287): 453–68.
Ehlers, Lars, Isa E. Hafalir, M. Bumin Yenmez, and Muhammed A. Yildirim. 2014. “School Choice with Controlled Choice Constraints: Hard Bounds Versus Soft Bounds.” *Journal of Economic Theory* 153 (September): 648–83.
Eissa, Nada, and Hilary W Hoynes. 2006. “Behavioral Responses to Taxes: Lessons from the EITC and Labor Supply.” *Tax Policy and the Economy* 20: 73–110.
Eissa, Nada, Henrik Jacobsen Kleven, and Claus Thustrup Kreiner. 2006. “Welfare Effects of Tax Reform and Labor Supply at the Intensive and Extensive Margins.” In *Tax Policy and Labor Market Performance*, edited by Jonas Agell and Peter Birch Sørensen, 147–86. The MIT Press.
Elango, Sneha, Jorge Luis Garcìa, James J. Heckman, and Andres Hojman. 2016. “Early Childhood Education.” In *Economics of Means-Tested Transfer Programs in the United States*, Moffitt, Robert A. Vol. II. University of Chicago Press.
Elbers, Benjamin, and Rob J. Gruijters. 2024. “Segplot: A New Method for Visualizing Patterns of Multi-Group Segregation.” *Research in Social Stratification and Mobility* 89 (February): 100860.
Eliaz, Kfir, and Ran Spiegler. 2020. “A Model of Competing Narratives.” *American Economic Review* 110 (12): 3786–816.
Ellwood, David T. 2000. “Anti-Poverty Policy for Families in the Next Century: From Welfare to Work—and Worries.” *Journal of Economic Perspectives* 14 (1): 187–98.
Engbom, Niklas, Enrica Detragiache, and Faezeh Raei. 2015. “The German Labor Market Reforms and Post-Unemployment Earnings.” Working Paper WP/15/162. Washington DC: International Monetary Fund.
Ermisch, John, Marco Francesconi, and David J. Pevalin. 2004. “Parental Partnership and Joblessness in Childhood and Their Influence on Young People’s Outcomes.” *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 167 (1): 69–101.
Esping-Andersen, G. 1990. *The Three Worlds of Welfare Capitalism*. Cambridge. Polity Press.
Européenne, Commission. 2015. “Investissement Social - Emploi, Affaires Sociales Et Inclusion - Commission Européenne.” 2015.
Evans, Gary W., and Kalee De France. 2022. “Childhood Poverty and Psychological Well-Being: The Mediating Role of Cumulative Risk Exposure.” *Development and Psychopathology* 34 (3): 911–21.
Evans, Patricia M. 2007. “(Not) Taking Account of Precarious Employment: Workfare Policies and Lone Mothers in Ontario and the UK.” *Social Policy & Administration* 41 (1): 29–49.
Eydoux, Anne, and Marie-Thérèse Letablier. 2009. “Familles Monoparentales Et Pauvreté En Europe : Quelles Réponses Politiques ?” *Politiques Sociales Et Familiales*, no. 98: 21.
Fack, Gabrielle, and Julien Grenet. 2010a. “When Do Better Schools Raise Housing Prices? Evidence from Paris Public and Private Schools.” *Journal of Public Economics* 94 (1-2): 59–77.
———. 2010b. “When Do Better Schools Raise Housing Prices? Evidence from Paris Public and Private Schools.” *Journal of Public Economics* 94 (1-2): 59–77.
Fagnani, Jeanne. 2010. “Childcare Policies in France: The Influence of Organizational Changes in the Workplace.” In *From Child Welfare to Child Well-Being: An International Perspective on Knowledge in the Service of Policy Making*, edited by Sheila B. Kamerman, Shelley Phipps, and Asher Ben-Arieh, 385–402. Children’s Well-Being: Indicators and Research. Dordrecht: Springer Netherlands.
Farfan-Portet, Maria-Isabel, Vincent Lorant, and Francesca Petrella. 2011. “Access to Childcare Services: The Role of Demand and Supply-Side Policies.” *Population Research and Policy Review* 30 (2): 165–83.
Fehr, Ernst, and Lorenz Goette. 2007. “Do Workers Work More If Wages Are High? Evidence from a Randomized

Field Experiment." *The American Economic Review* 97 (1): 298–317.
Felfe, Christina, and Rafael Lalive. 2018. "Does Early Child Care Affect Children's Development?" *Journal of Public Economics* 159 (March): 33–53.
Felfe, Christina, Natalia Nollenberger, and Núria Rodríguez-Planas. 2015. "Can't Buy Mommy's Love? Universal Childcare and Children's Long-Term Cognitive Development." *Journal of Population Economics* 28 (2): 393–422.
Fenech, Marianne, and Jennifer Skattebol. 2019. "Supporting the Inclusion of Low-Income Families in Early Childhood Education: An Exploration of Approaches Through a Social Justice Lens." *International Journal of Inclusive Education*, April, 1–19.
Fernald, Lia C H, Elizabeth Prado, Patricia Kariger, and Abbie Raikes. 2017. "A Toolkit for Measuring Early Childhood Development in Low- and Middle-Income Countries." Prepared for the Strategic Impact Evaluation Fund, the World Bank. Washington DC: World Bank.
Ferrari, Giulia, Carole Bonnet, and Anne Solaz. 2019. "'Will the One Who Keeps the Children Keep the House?' Residential Mobility After Divorce by Parenthood Status and Custody Arrangements in France." *Demographic Research* 40 (February): 359–94.
Filges, Trine, G Smedslund, Anne-Sophie Due Knudsen, and Anne-Marie Klint Jørgensen. 2015. "Active Labour Market Programme (ALMP) Participation for Unemployment Insurance Recipients." The Campbell Collaboration.
Finkelstein, Amy, and Matthew J Notowidigdo. 2019. "Take-Up and Targeting: Experimental Evidence from SNAP*." *The Quarterly Journal of Economics* 134 (3): 1505–56.
Firpo, Sergio. 2007. "Efficient Semiparametric Estimation of Quantile Treatment Effects." *Econometrica* 75 (1): 259–76.
Fischer, Ronald A. 1935. *The Design of Experiments*. HAFNER PRESS. London: COLLIER MACMILLAN PUBLISERS.
Fisher, Tracy, and Ellen Reese. 2011. "The Punitive Turn in Social Policies: Critical Race Feminist Reflections on the USA, Great Britain, and Beyond." *Critical Sociology* 37 (2): 225–36.
Fleche, Sarah, Anthony Lepinteur, and Nattavudh Powdthavee. 2018a. "Gender Norms and Relative Working Hours: Why Do Women Suffer More Than Men from Working Longer Hours Than Their Partners?" In *AEA Papers and Proceedings*, 108:163–68. American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203.
———. 2018b. "Gender Norms and Relative Working Hours: Why Do Women Suffer More Than Men from Working Longer Hours Than Their Partners?" *AEA Papers and Proceedings* 108 (May): 163–68.
Flèche, Sarah, Anthony Lepinteur, and Nattavudh Powdthavee. 2020. "Gender Norms, Fairness and Relative Working Hours Within Households." *Labour Economics* 65 (August): 101866.
Fleurbaey, Marc, and François Maniquet. 2018. "Optimal Income Taxation Theory and Principles of Fairness." *Journal of Economic Literature* 56 (3): 1029–79.
Foote, Christopher L., and Christopher F. Goetz. 2008. "The Impact of Legalized Abortion on Crime: Comment*." *The Quarterly Journal of Economics* 123 (1): 407–23.
FORS. 2020. "Mission d'évaluation Du Projet d'investissement Social RELIANCE Rapport Final." Rapport d'évaluation. Conseil Départemental Meurthe-Et-Moselle.
Fortin, N. M. 2005. "Gender Role Attitudes and the Labour-market Outcomes of Women Across OECD Countries." *Oxford Review of Economic Policy* 21 (3): 416–38.
Fortin, Nicole, Thomas Lemieux, and Sergio Firpo. 2011. "Decomposition Methods in Economics." In *Handbook of Labor Economics*, 4:1–102. Elsevier.
Foster, Carly Hayden. 2008. "The Welfare Queen: Race, Gender, Class, and Public Opinion." *Race, Gender & Class* 15 (3/4): 162–79.
Fougère, Denis, and Arthur Heim. 2019. "L'évaluation Socioéconomique de l'investissement Social." Rapport 2019-06. France Stratégie.
Francesconi, M., and J. J. Heckman. 2016. "Child Development and Parental Investment: Introduction." *The Economic Journal* 126 (596, 596): F1–27.
Freedman, Seth M., Alex Hollingsworth, Kosali I. Simon, Coady Wing, and Madeline Yozwiak. 2023. "Designing Difference in Difference Studies With Staggered Treatment Adoption: Key Concepts and Practical Guidelines." Working Paper. Working Paper Series. November 2023.
Frémeaux, Nicolas, and Marion Leturcq. 2022. "Wealth Accumulation and the Gender Wealth Gap Across Couples'

Legal Statuses and Matrimonial Property Regimes in France.” *European Journal of Population* 38 (4): 643–79.
Friedrichsen, Jana, Tobias König, and Renke Schmacker. 2018. “Social Image Concerns and Welfare Take-up.” *Journal of Public Economics* 168 (December): 174–92.
Frisch, Ragnar. 1930. “A Dynamic Approach to Economic Theory.” In, 246. Yale University.
Frölich, Markus, and Blaise Melly. 2013. “Identification of Treatment Effects on the Treated with One-Sided Non-Compliance.” *Econometric Reviews* 32 (3): 384–414.
Gale, D., and L. S. Shapley. 1962. “College Admissions and the Stability of Marriage.” *The American Mathematical Monthly* 69 (1): 9.
Galitzine, Alexandra, and Arthur Heim. 2024. “Tax Burden on the Poor: Single Mothers’ Optimisation Behaviours Following an Experimental Activation Programme in France.” In *Social Investment & The Changing Face of Poverty: Essays on the Design and Evaluation of Family and Social Policies in France*, Paris School of Economics. Paris (France): PhD Thesis.
Gambaro, Ludovica, Kitty Stewart, and Jane Waldfogel. 2015. “A Question of Quality: Do Children from Disadvantaged Backgrounds Receive Lower Quality Early Childhood Education and Care?” *British Educational Research Journal* 41 (4): 553–74.
Gansen, Heidi M. 2017. “Reproducing (and Disrupting) Heteronormativity: Gendered Sexual Socialization in Preschool Classrooms.” *Sociology of Education* 90 (3): 255–72.
Garbinti, Bertrand, Jonathan Goupille-Lebret, Mathilde Munoz, Stefanie Stantcheva, and Gabriel Zucman. 2023. “Tax Design, Information, and Elasticities: Evidence From the French Wealth Tax.”
Garritzmann, Julian L., Silja Häusermann, and Bruno Palier, eds. 2022. *The World Politics of Social Investment: Volume I: Welfare States in the Knowledge Economy*. 1st ed. Oxford University PressNew York.
Gazier, Bernard. 2019. “Opportunities or Tensions: Assessing French Labour Market Reforms from 2012 to 2018.” *International Journal of Comparative Labour Law and Industrial Relations* 35 (September): 331–54.
Gelber, Alexander M., Damon Jones, and Daniel W. Sacks. 2020. “Estimating Adjustment Frictions Using Nonlinear Budget Sets: Method and Evidence from the Earnings Test.” *American Economic Journal: Applied Economics* 12 (1): 1–31.
Gelber, Alexander M., and Joshua W. Mitchell. 2012. “Taxes and Time Allocation: Evidence from Single Women and Men.” *The Review of Economic Studies* 79 (3): 863–97.
Gendre, Alexandra de, Stefanie Schurer, and Angela Zhang. 2022. “Two Decades of Welfare Reforms in Australia: How Did They Affect Single Mothers and Their Children?” *SSRN Electronic Journal*.
Gennetian, Lisa, Greg Duncan, Nathan Fox, Katherine Magnuson, Sarah Halpern-Meekin, Kimberly Noble, and Hirokazu Yoshikawa. 2022. “Unconditional Cash and Family Investments in Infants: Evidence from a Large-Scale Cash Transfer Experiment in the U.S.” w30379. Cambridge, MA: National Bureau of Economic Research.
Gestel, Nicolette van, and Jean□Michel Herbillon. 2007. “Changing Modes of Governance in Activation Policies in France and The Netherlands: Common Path or Countermodel?” Edited by Rik van Berkel. *International Journal of Sociology and Social Policy* 27 (7/8): 324–33.
Ghysels, Joris, and Wim Van Lancker. 2011. “The Unequal Benefits of Activation: An Analysis of the Social Distribution of Family Policy Among Families with Young Children.” *Journal of European Social Policy* 21 (5): 472–85.
Gibson, Marcia, Hilary Thomson, Kasia Banas, Vittoria Lutje, Martin J McKee, Susan P Martin, Candida Fenton, Clare Bambra, and Lyndal Bond. 2018. “Welfare-to-Work Interventions and Their Effects on the Mental and Physical Health of Lone Parents and Their Children.” Edited by Cochrane Public Health Group. *Cochrane Database of Systematic Reviews* 2018 (2).
Gobillon, Laurent, Dominique Meurs, and Sébastien Roux. 2015. “Estimating Gender Differences in Access to Jobs.” *Journal of Labor Economics* 33 (2): 317–63.
Goldin, C., and Lawrence F Katz. 2008. *The Race Between Education and Technology*. Cambridge, MA: Harvard University Press.
Goldin, Claudia. 2006-May. “The Quiet Revolution That Transformed Women’s Employment, Education, and Family.” *AEA Papers and Proceedings* 96 (2): 21.
Goldin, Claudia, and Lawrence F Katz. 2002. “The Power of the Pill: Oral Contraceptives and Women’s Career and Marriage Decisions.” *Journal of Political Economy* 110 (4): 730–70.
Gomajee, Ramchandar, Fabienne El-Khoury, Sylvana Côté, Judith van der Waerden, Laura Pryor, and Maria Melchior. 2018. “Early Childcare Type Predicts Children’s Emotional and Behavioural Trajectories into Middle Childhood. Data from the EDEN Mother–Child Cohort Study.” *Journal of Epidemiology and Community*

Health, October, jech-2017-210393.
Gomajee, Ramchandar, Fabienne El-Khoury, Judith van der Waerden, Laura Pryor, and Maria Melchior. 2017. "Early Life Childcare and Later Behavioral Difficulties: A Causal Relationship? Data from the French EDEN Study." *Journal of Economic Behavior & Organization*, December, S0167268117303372.
Gonalons-Pons, Pilar, and Christine R. Schwartz. 2017. "Trends in Economic Homogamy: Changes in Assortative Mating or the Division of Labor in Marriage?" *Demography* 54 (3): 985–1005.
Gong, Yifan, Ralph Stinebrickner, and Todd Stinebrickner. 2022. "Marriage, Children, and Labor Supply: Beliefs and Outcomes." *Journal of Econometrics*, Annals Issue: Subjective Expectations & Probabilities in Economics., 231 (1): 148–64.
Gonzalez, Libertad, Luis Guirola, and Blanca Zapater. 2022. "Partisan Abortions." March 2022.
Goodman-Bacon, Andrew. 2021. "Difference-in-Differences with Variation in Treatment Timing." *Journal of Econometrics*, June, S0304407621001445.
Gorey, Kevin. 2009. "Welfare-to-Work Programs in America, 1980 to 2005: Meta-Analytic Evidence of the Importance of Job and Child Care Availability." *Journal of Policy Practice* 8 (4): 265–81.
Goussé, Marion, Nicolas Jacquemet, and Jean-Marc Robin. 2017. "Marriage, Labor Supply, and Home Production." *Econometrica* 85 (6): 1873–1919.
Goux, Dominique, and Eric Maurin. 2010. "Public School Availability for Two-Year Olds and Mothers' Labour Supply." *Labour Economics* 17 (6): 951–62.
Greenwood, Jeremy, Ananth Seshadri, and Mehmet Yorukoglu. 2005. "Engines of Liberation." *The Review of Economic Studies* 72 (1): 109–33.
Gregg, Paul, Susan Harkness, and Sarah Smith. 2009. "Welfare Reform and Lone Parents in the UK." *The Economic Journal* 119 (535): F38–65.
Grenet, Julien, YingHua He, and Dorothea Kübler. 2022. "Preference Discovery in University Admissions: The Case for Dynamic Multioffer Mechanisms." *Journal of Political Economy* 130 (6): 1427–76.
Grey, E. B., L. Atkinson, A. Chater, A. Gahagan, A. Tran, and F. B. Gillison. 2022. "A Systematic Review of the Evidence on the Effect of Parental Communication about Health and Health Behaviours on Children's Health and Wellbeing." *Preventive Medicine* 159 (June): 107043.
Grobon, Sébastien, Lidia Panico, and Anne Solaz. 2019. "Inégalités Socioéconomiques Dans Le Développement Langagier Et Moteur Des Enfants à 2 Ans." *Bulletin Epidemiologique Hebdomadaire*, no. 1: 8.
Grogger, Jeffrey. 2003. "The Effects of Time Limits, the EITC, and Other Policy Changes on Welfare Use, Work, and Income Among Female-Headed Families." *Review of Economics and Statistics* 85 (2): 394–408.
Groot, Wim, and Henriëtte Maassen van den Brink. 2000. "Overeducation in the Labor Market: A Meta-Analysis." *Economics of Education Review* 19 (2): 149–58.
Gruber, Jonathan, Phillip Levine, and Douglas Staiger. 1999. "Abortion Legalization and Child Living Circumstances: Who Is the 'Marginal Child'?*." *The Quarterly Journal of Economics* 114 (1): 263–91.
Guillen, Pablo, and Rustamdjan Hakimov. 2018. "The Effectiveness of Top-down Advice in Strategy-Proof Mechanisms: A Field Experiment." *European Economic Review* 101 (January): 505–11.
Gupta, Datta, Nabanita Jessen, and Jonas. 2023. "Maternity Leave Versus Early Childcare—What Are the Long-Term Consequences for Children?" *IZA World of Labor*, May.
Gupta, Nabanita Datta, and Marianne Simonsen. 2010. "Non-Cognitive Child Outcomes and Universal High Quality Child Care." *Journal of Public Economics* 94 (1-2): 30–43.
Gurgand, Marc, and David N. Margolis. 2008. "Does Work Pay in France? Monetary Incentives, Hours Constraints, and the Guaranteed Minimum Income." *Journal of Public Economics* 92 (7): 1669–97.
Hakovirta, Mia, Johanna Kallio, and Milla Salin. 2021. "Is It Possible for Single Parents to Successfully Raise Children? Multilevel Analysis of Attitudes Toward Single Parents in 22 Welfare States." *Journal of Comparative Family Studies* 52 (1): 117–44.
Hannafi, Cyrine, Céline Marc, Rémi Le Gall, and Laure Omalek. 2022. "Regularly Measuring the Non-Take-up of the RSA and the Employment Bonus: Method and Results." 92. Paris: DREES.
Hargreaves, Alec G. 2015. "Empty Promises?: Public Policy Against Racial and Ethnic Discrimination in France." *French Politics, Culture & Society* 33 (3): 95–115.
Hartley, Robert Paul, Carlos Lamarche, and James P. Ziliak. 2022. "Welfare Reform and the Intergenerational Transmission of Dependence." *Journal of Political Economy* 130 (3): 523–65.
Hassidim, Avinatan, Assaf Romm, and Ran I. Shorrer. 2018. "'Strategic' Behavior in a Strategy-Proof Environment." SSRN Scholarly Paper. Rochester, NY. February 22, 2018.

Häusermann, Silja. 2018. “The Multidimensional Politics of Social Investment in Conservative Welfare Regimes: Family Policy Reform Between Social Transfers and Social Investment.” *Journal of European Public Policy* 25 (6): 862–77.

Havnes, Tarjei, and Magne Mogstad. 2015. “Is Universal Child Care Leveling the Playing Field?” *Journal of Public Economics* 127 (July): 100–114.

Heckman, James. 1990. “Varieties of Selection Bias.” *American Economic Review* 80 (2): 313–18.

Heckman, James J., Robert J. Lalonde, and Jeffrey A. Smith. 1999. “The Economics and Econometrics of Active Labor Market Programs.” In *Handbook of Labor Economics*, 3:1865–2097. Elsevier.

Heckman, James J., and Rodrigo Pinto. 2022a. “The Econometric Model for Causal Policy Analysis.” *Annual Review of Economics* 14 (1): 893–923.

———. 2022b. “The Econometric Model for Causal Policy Analysis.” *Annual Review of Economics* 14 (1): 893–923.

Heckman, James J., and Jeffrey A. Smith. 1999. “The Pre-Programme Earnings Dip and the Determinants of Participation in a Social Programme. Implications for Simple Programme Evaluation Strategies.” *The Economic Journal* 109 (457): 313–48.

Heckman, James, Hidehiko Ichimura, Jeffrey Smith, and Petra Todd. 1998. “Characterizing Selection Bias Using Experimental Data.” *Econometrica* 66 (5): 1017–98.

Heim, Arthur. 2017. “Peut-on Estimer Le Rendement de l’investissement Social ?” Document de travail 2017-02. France Stratégie.

———. 2024. “Welfare-to-What ? Experimental Evaluation of an Activation Programme for Single Mothers in Poverty in France.” In *Social Investment & The Changing Face of Poverty: Essays on the Design and Evaluation of Family and Social Policies in France*, Paris School of Economics. 2. Paris (France): PhD Thesis.

Hemerijck, A. C. 2014. “Social Investment ‘Stocks,’ ‘Flows’ and ‘Buffers’.” *Rivista Delle Politiche Sociali* 1 (1): 9–26.

Hemerijck, Anton. 2018. “Social Investment as a Policy Paradigm.” *Journal of European Public Policy* 25 (6): 810–27.

Hemerijck, Anton, and Robin Huguenot-Noël. 2022. *Resilient Welfare States in the European Union*. Comparative Political Economy. Newcastle upon Tyne: Agenda Publishing.

Henderson, Heath, and Lendie Follett. 2022. “Targeting Social Safety Net Programs on Human Capabilities.” *World Development* 151 (March): 105741.

Herbst, Chris M. 2011. “The Impact of the Earned Income Tax Credit on Marriage and Divorce: Evidence from Flow Data.” *Population Research and Policy Review* 30 (1).

Herbst-Debby, Anat. 2022. “(De)legitimization of Single Mothers’ Welfare Rights: United States, Britain and Israel.” *Journal of European Social Policy* 32 (3): 302–16.

Herd, Pamela, Hilary Hoynes, Jamila Michener, and Donald Moynihan. 2023. “Introduction: Administrative Burden as a Mechanism of Inequality in Policy Implementation.” *RSF: The Russell Sage Foundation Journal of the Social Sciences* 9 (5): 1–30.

Herke, Boglárka. 2021. “Investigating the Welfare Deservingness of Single Mothers: Public Image and Deservingness Perceptions in Hungary.” *East European Politics and Societies: And Cultures* 35 (3): 613–37.

Herman, Elisa. 2017. “Politiques Locales de La Petite Enfance : Les Enjeux de Qualification de La Demande de Place En Établissement d’accueil Du Jeune Enfant.” *Revue Francaise Des Affaires Sociales*, no. 2 (July): 41–61.

Hermes, Henning, Marina Krauß, Philipp Lergetporer, Frauke Peter, and Simon Wiederhold. 2022. “Early Child Care and Labor Supply of Lower-Ses Mothers: A Randomized Controlled Trial.” *SSRN Electronic Journal*.

Hermes, Henning, Philipp Lergetporer, Frauke Peter, and Simon Wiederhold. 2021. “Behavioral Barriers and the Socioeconomic Gap in Child Care Enrollment.” *SSRN Electronic Journal*.

Herzog, Sabrina, and Thilo Klein. 2018. “Matching Practices for Childcare - Germany.” MiP country profile.

Hewlett, Nick, and Raymond Kuhn. 2022. “Reflections on the 2022 Elections in France.” *Modern & Contemporary France* 30 (4): 393–409.

Hirano, Kiesukie, Guido W Imbens, and Geert Ridder. 2003. “Efficient Estimation of Average Treatment Effects Using the Estimated Propensity Score.” *Econometrica* 71 (4): 1161–89.

Hirschman, Daniel, and Elizabeth Popp Berman. 2014. “Do Economists Make Policies? On the Political Effects of Economics1.” *Socio-Economic Review* 12 (4): 779–811.

Hoffmann, Florian, David S. Lee, and Thomas Lemieux. 2020. “Growing Income Inequality in the United States and Other Advanced Economies.” *Journal of Economic Perspectives* 34 (4): 52–78.

Hohmeyer, Katrin, and Joachim Wolff. 2018. “Of Carrots and Sticks: The Effect of Workfare Announcements on the Job Search Behaviour and Reservation Wage of Welfare Recipients.” *Journal for Labour Market Research* 52 (1): 11.

Holloway, Sarah L. 1998. “Local Childcare Cultures: Moral Geographies of Mothering and the Social Organisation of Pre-School Education.” *Gender, Place & Culture* 5 (1): 29–53.

Hothorn, Torsten, Frank Bretz, and Peter Westfall. 2008. “Simultaneous Inference in General Parametric Models.” *Biometrical Journal* 50 (3): 346–63.

Howard-Jones, P. A., E. V. Washbrook, and S. Meadows. 2012. “The Timing of Educational Investment: A Neuroscientific Perspective.” *Developmental Cognitive Neuroscience* 2 (February): S18–29.

Hoynes, Hilary, Doug Miller, and David Simon. 2015. “Income, the Earned Income Tax Credit, and Infant Health.” *American Economic Journal: Economic Policy* 7 (1): 172–211.

Huber, Katrin, and Erwin Winkler. 2016. “All We Need Is Love? Trade-Adjustment, Inequality, and the Role of the Partner.” SSRN Scholarly Paper. Rochester, NY. October 10, 2016.

Huber, Martin, and Kaspar Wüthrich. 2019. “Local Average and Quantile Treatment Effects Under Endogeneity: A Review.” *Journal of Econometric Methods* 8 (1).

Hufe, Paul, Ravi Kanbur, and Andreas Peichl. 2020. “Measuring Unfair Inequality: Reconciling Equality of Opportunity and Freedom from Poverty.” Discussion paper 232. Ludwig-Maximilians-Universität München und Humboldt-Universität zu Berlin: Collaborative Research Center Transregio 190 - Rationality and Competition.

Huizen, Thomas van, and Janneke Plantenga. 2018. “Do Children Benefit from Universal Early Childhood Education and Care? A Meta-Analysis of Evidence from Natural Experiments.” *Economics of Education Review* 66 (October): 206–22.

Hutchens, Robert M. 1991. “Segregation Curves, Lorenz Curves, and Inequality in the Distribution of People Across Occupations.” *Mathematical Social Sciences* 21 (1): 31–51.

Hvidberg, Kristoffer B, Claus T Kreiner, and Stefanie Stantcheva. 2022. “Social Positions and Fairness Views on Inequality,” December.

———. 2023. “Social Positions and Fairness Views on Inequality.” *The Review of Economic Studies* 90 (6): 3083–3118.

Ichino, Andrea, Margherita Fort, and Giulio Zanella. 2019. “Cognitive and Non-Cognitive Costs of Daycare 0-2 for Children in Advantaged Families.” *Journal of Political Economy*, April, 704075.

Imbens, Guido W. 2020. “Potential Outcome and Directed Acyclic Graph Approaches to Causality: Relevance for Empirical Practice in Economics.” *Journal of Economic Literature* 58 (4): 1129–79.

INSEE. 2021a. “Les Familles En 2020 : 25 % de Familles Monoparentales, 21 % de Familles Nombreuses - Insee Focus - 249.” 249.

———. 2021b. “Neuf Mois Après Le Premier Confinement, Une Baisse Plus Marquée Des Naissances Dans Les Territoires Fortement Touchés Par l’épidémie.” *INSEE Références*.

———. 2023. “Dossier Complet − Arrondissement de Nancy (543) | Insee.” January 2023.

Ishii, Karine, Per Yann LE Floc’H, Adrien Massebieau, and Baptiste Royer. 2023. “Les Inégalités d’accès Aux Crèches Et Leurs Enjeux Économiques.” *Direction Générale Du Trésor*, Trésor-eco, 322 (January).

Jackson, Margot I., Kathleen Kiernan, and Sara McLanahan. 2017. “Maternal Education, Changing Family Circumstances, and Children’s Skill Development in the United States and UK.” *The ANNALS of the American Academy of Political and Social Science* 674 (1): 59–84.

Jacquey-Vazquez, Benedicte. 2017. “L’accompagnement Social.” RAPPORT IGAS 2017-105K.

Jaehrling, K., T. Kalina, and L. Mesaros. 2015. “A Paradox of Activation Strategies: Why Increasing Labour Market Participation Among Single Mothers Failed to Bring Down Poverty Rates.” *Social Politics: International Studies in Gender, State & Society* 22 (1): 86–110.

Jaroszewicz, Ania, Jon Jachimowicz, Oliver Hauser, and Julian Jamison. 2022. “How Effective Is (More) Money? Randomizing Unconditional Cash Transfer Amounts in the US.” *SSRN Electronic Journal*.

Jarrett, Robin L. 1996. “Welfare Stigma Among Low-Income, African American Single Mothers.” *Family Relations* 45 (4): 368–74.

Jenson, Jane. 2010. “Diffusing Ideas for After Neoliberalism: The Social Investment Perspective in Europe and Latin America.” *Global Social Policy: An Interdisciplinary Journal of Public Policy and Social Development* 10 (1): 59–84.

Johnsen, Julian Vedeler, and Katrine Holm Reiso. 2020. “Economic Effects of Workfare Reforms for Single Mothers: Benefit Substitution and Labour Supply Responses*." *The Scandinavian Journal of Economics* 122 (2): 494–523.

Johnson, Rucker C, and Robert F Schoeni. 2011. "The Influence of Early-Life Events on Human Capital, Health Status, and Labor Market Outcomes Over the Life Course." *The B.E. Journal of Economic Analysis & Policy* 11 (3).

Johnson, Rucker, and C. Kirabo Jackson. 2017. "Reducing Inequality Through Dynamic Complementarity: Evidence from Head Start and Public School Spending." w23489. Cambridge, MA: National Bureau of Economic Research.

Jones, Damon, and Ioana Marinescu. 2022. "The Labor Market Impacts of Universal and Permanent Cash Transfers: Evidence from the Alaska Permanent Fund." *American Economic Journal: Economic Policy* 14 (2): 315–40.

Joyce, Ted. 2001. "Did Legalized Abortion Lower Crime ?" NBER Working Paper Series. Cambridge, MA. June 2001.

———. 2009. "A Simple Test of Abortion and Crime." *The Review of Economics and Statistics* 91 (1): 112–23.

Juhn, Chinhui, and Kristin McCue. 2017. "Specialization Then and Now: Marriage, Children, and the Gender Earnings Gap Across Cohorts." *Journal of Economic Perspectives* 31 (1): 183–204.

Kamada, Yuichiro, and Fuhito Kojima. 2019. "Fair Matching Under Constraints: Theory and Applications." *R&R in Review of Economic Studies*, 59.

———. 2023. "Fair Matching Under Constraints: Theory and Applications." *The Review of Economic Studies*, April.

Kandil, Lamia, Hélène Périvier, and Camille Richou. 2021. "Partager Les Tâches Domestiques ? La Division Du Travail Dans Les Couples Selon Le Type d'union En France, 1985-2009:" *Population* Vol. 76 (1): 155–91.

Kasy, Maximilian. 2023. "The Political Economy of AI: Towards Democratic Control of the Means of Prediction." Preprint. SocArXiv.

Kasy, Maximilian, and Rediet Abebe. 2021. "Fairness, Equality, and Power in Algorithmic Decision-Making." In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 576–86. Virtual Event Canada: ACM.

Kasy, Maximilian, and Lukas Lehner. 2023. "Employing the Unemployed of Marienthal: Evaluation of a Guaranteed Job Program."

Katz, Lawrence F., and David H. Autor. 1999. "Changes in the Wage Structure and Earnings Inequality." In *Handbook of Labor Economics*, 3:1463–1555. Elsevier.

Kendig, Sarah M., and Suzanne M. Bianchi. 2008. "Single, Cohabitating, and Married Mothers' Time With Children." *Journal of Marriage and Family* 70 (5): 1228–40.

Kennes, John, Daniel Monte, and Norovsambuu Tumennasan. 2014. "The Day Care Assignment: A Dynamic Matching Problem." *American Economic Journal: Microeconomics* 6 (4): 362–406.

Kerr, N. L. 1998. "HARKing: Hypothesizing After the Results Are Known." *Personality and Social Psychology Review: An Official Journal of the Society for Personality and Social Psychology, Inc* 2 (3): 196–217.

Kesten, Onur. 2010. "School Choice with Consent *." *Quarterly Journal of Economics* 125 (3): 1297–1348.

Kholoptseva, Evgenia. 2016. "Effects of Center-Based Early Childhood Education Programs on Children's Language, Literacy, and Math Skills: A Comprehensive Meta-Analysis," May.

Kimmel, Jean. 1998. "Child Care Costs as a Barrier to Employment for Single and Married Mothers." *The Review of Economics and Statistics* 80 (2): 287–99.

Kleven, Henrik. 2019. "The EITC and the Extensive Margin: A Reappraisal." w26405. Cambridge, MA: National Bureau of Economic Research.

Kleven, Henrik Jacobsen. 2016. "Bunching." *Annual Review of Economics* 8 (1): 435–64.

———. 2023. "The EITC and the Extensive Margin: A Reappraisal." w26405. Cambridge, MA: National Bureau of Economic Research.

Kleven, Henrik Jacobsen, and Wojciech Kopczuk. 2011. "Transfer Program Complexity and the Take-Up of Social Benefits." *American Economic Journal: Economic Policy* 3 (1): 54–90.

Kleven, Henrik J., and Mazhar Waseem. 2013. "Using Notches to Uncover Optimization Frictions and Structural Elasticities: Theory and Evidence from Pakistan*." *The Quarterly Journal of Economics* 128 (2): 669–723.

Kleven, Henrik, Camille Landais, and Gabriel Leite-Mariante. 2024. "The Child Penalty Atlas." Working Paper 31649. NBER.

Kleven, Henrik, Camille Landais, Johanna Posch, Andreas Steinhauer, and Josef Zweimüller. 2023. "Do Family Policies Reduce Gender Inequality? Evidence from 60 Years of Policy Experimentation." *American Economic*

Journal: Economic Policy Forthcoming.

Kleven, Henrik, Camille Landais, and Jakob Egholt Søgaard. 2021. "Does Biology Drive Child Penalties? Evidence from Biological and Adoptive Families." *American Economic Review: Insights* 3 (2): 183–98.

Kleven, Henrik, Camillie Landais, and Jakob Egholt Søgaard. 2019. "Children and Gender Inequality: Evidence from Denmark." *American Economic Journal: Applied Economics* 11 (4): 181–209.

Kline, Patrick, and Melissa Tartari. 2016. "Bounding the Labor Supply Responses to a Randomized Welfare Experiment: A Revealed Preference Approach." *American Economic Review* 106 (4): 972–1014.

Knapp, Andrew. 2022. "France's Party System in 2022." *Modern & Contemporary France* 30 (4): 495–515.

Knijn, Trudie, Claude Martin, and Jane Millar. 2007. "Activation as a Common Framework for Social Policies Towards Lone Parents." *Social Policy & Administration* 41 (6): 638–52.

Ko, Wonsik, and Robert A Moffitt. 2022. "Take-up of Social Benefits." Discussion Paper. 2022.

Kominers, Scott Duke, Alexander Teytelboym, and Vincent P Crawford. 2017. "An Invitation to Market Design." *Oxford Review of Economic Policy* 33 (4): 541–71.

Kostøl, Andreas R., and Andreas S. Myhre. 2021. "Labor Supply Responses to Learning the Tax and Benefit Schedule." *American Economic Review* 111 (11): 3733–66.

Kottelenberg, Michael J., and Steven F. Lehrer. 2016. "Targeted or Universal Coverage? Assessing Heterogeneity in the Effects of Universal Child Care." *Journal of Labor Economics* 35 (3): 609–53.

Kovacic, Matija, and Cristina Elisa Orso. 2023. "Who's Afraid of Immigration? The Effect of Economic Preferences on Tolerance." *Journal of Population Economics* 36 (3): 1901–40.

Kramarz, Francis, Elio Nimier-David, and Thomas Delemotte. 2022. "Inequality and Earnings Dynamics in France: National Policies and Local Consequences." *Quantitative Economics* 13 (4): 1527–91.

Kreiner, Claus Thustrup, and Torben Tranaes. 2005. "Optimal Workfare with Voluntary and Involuntary Unemployment*." *Scandinavian Journal of Economics* 107 (3): 459–74.

Krueger, Alan B. 2008. "Are We Having More Fun Yet? Categorizing and Evaluating Changes in Time Allocation." *Brookings Papers on Economic Activity* 2007 (2): 193–215.

Kuitto, Kati. 2016. "From Social Security to Social Investment? Compensating and Social Investment Welfare Policies in a Life-Course Perspective." *Journal of European Social Policy* 26 (5): 442–59.

Kulic, Nevena, Jan Skopek, Moris Triventi, and Hans-Peter Blossfeld. 2019. "Social Background and Children's Cognitive Skills: The Role of Early Childhood Education and Care in a Cross-National Perspective." *Annual Review of Sociology* 45 (1): 557–79.

Kuziemko, Ilyaba, Michael I. Norton, Emmanuel Saez, and Stefanie Stantcheva. 2015. "How Elastic Are Preferences for Redistribution? Evidence from Randomized Survey Experiments." *The American Economic Review* 105 (4): 1478–508.

Kvist, Jon. 2015. "A Framework for Social Investment Strategies: Integrating Generational, Life Course and Gender Perspectives in the EU Social Investment Strategy." *Comparative European Politics* 13 (1): 131–49.

Lacouette Fougère, Clément, and Pierre Lascoumes. 2013. "L'évaluation : Un Marronnier de l'action Gouvernementale ?" *Revue Française d'administration Publique* 148 (4): 859–75.

Laithier, Élisabeth. 2018. "Attribution Des Places En Crèche." Vade Meccum. Paris: Association des Maires de France.

Lalonde, Robert J. 1986. "Evaluating the Econometric Evaluations of Training Programs with Experimental Data." *The American Economic Review* 76 (4): 604–20.

Lancker, Wim Van, and Joris Ghysels. 2012. "Who Benefits? The Social Distribution of Subsidized Childcare in Sweden and Flanders." *Acta Sociologica* 55 (2): 125–42.

Laporte, Claire, Arnaud Crépin, and Damien Hilairet. 2019. "Attentes, Besoins Et Contraintes Des Parents En Matière de Conciliation Vie Familiale Et Vie Professionnelle." Dossier d'étude 208-2019. Caisse nationale des allocations familiales.

Latzman, Natasha E., Cecilia Casanueva, Julia Brinton, and Valerie L. Forman-Hoffman. 2019. "The Promotion of Well-Being Among Children Exposed to Intimate Partner Violence: A Systematic Review of Interventions." *Campbell Systematic Reviews* 15 (3): e1049.

Lavoie, Kévin, and Marie-Christine Saint-Jacques. 2020. "Lovers for a Time, Mothers for Life: Ecosystemic Analysis of Blended Family Experiences of Lesbian Mothers and Stepmothers." *Child & Family Social Work* 25 (4): 946–54.

Lawrence, Elizabeth, Elisabeth Root, and Stefanie Mollborn. 2015. "Residential Mobility in Early Childhood: Household and Neighborhood Characteristics of Movers and Non-Movers." *Demographic Research* 33 (Novem-

ber): 939–50.
Le Barbanchon, Thomas, Roland Rathelot, and Alexandra Roulet. 2020. “Gender Differences in Job Search: Trading Off COMMUTE AGAINST WAGE*.” *The Quarterly Journal of Economics* 136 (1): 381–426.
Le Forner, Hélène. 2020. “Age at Parents’ Separation and Achievement: Evidence from France Using a Sibling Approach.” *Annals of Economics and Statistics*, no. 138: 107–63.
Le Pape, Marie Clémence, and Clémence Helfter, eds. 2023. *Les Familles Monoparentales: Condition de Vie, Vécu Et Action Publique*. États Des Savoirs. Paris: La Documentation française.
Lefranc, Arnaud. 2003. “Labor Market Dynamics and Wage Losses of Displaced Workers in France and the United States.” *SSRN Electronic Journal*.
Leruth, Benjamin. 2017. “France at a Crossroads: Societal Challenges to the Welfare State During Nicolas Sarkozy’s and François Hollande’s Presidential Terms.” In *After Austerity: Welfare State Transformation in Europe After the Great Recession*, edited by Peter Taylor-Gooby, Benjamin Leruth, and Heejung Chung, 0. Oxford University Press.
Li, Shengwu. 2017a. “Ethics and Market Design.” *Oxford Review of Economic Policy* 33 (4): 705–20.
———. 2017b. “Obviously Strategy-Proof Mechanisms.” *American Economic Review* 107 (11): 3257–87.
Liang, Kung-Yee, and Scott L. Zeger. 1986. “Longitudinal Data Analysis Using Generalized Linear Models.” *Biometrika* 73 (1): 13–22.
Liebman, Jeffrey B., and Richard Zeckhauser. 2004. “Schmeduling.” NBER.
Lin, Winston. 2013. “Agnostic Notes on Regression Adjustments to Experimental Data: Reexamining Freedman’s Critique.” *The Annals of Applied Statistics* 7 (1): 295–318.
Linos, Elizabeth, Allen Prohofsky, Aparna Ramesh, Jesse Rothstein, and Matt Unrath. 2020. “Can Nudges Increase Take-up of the EITC?”
Lippmann, Quentin, Alexandre Georgieff, and Claudia Senik. 2020. “Undoing Gender with Institutions: Lessons from the German Division and Reunification.” *The Economic Journal* 130 (629): 1445–70.
Locks, Gedeão, and Josselin Thuilliez. 2023. “The Impact of Minimum Income on Homelessness: Evidence from France.” *Journal of Urban Economics* 135 (May): 103547.
Løken, Katrine V., Kjell Erik Lommerud, and Katrine Holm Reiso. 2018. “Single Mothers and Their Children: Evaluating a Work-Encouraging Welfare Reform.” *Journal of Public Economics* 167 (November): 1–20.
Lu, Li, Benjamin Contrand, Marion Dupuy, Leila Ramiz, Catherine Sztal-Kutas, and Emmanuel Lagarde. 2022. “Mental and Physical Health Among the French Population Before and During the First and Second COVID-19 Lockdowns: Latent Class Trajectory Analyses Using Longitudinal Data.” *Journal of Affective Disorders* 309 (July): 95–104.
Lubotsky, Darren, and Robert Kaestner. 2016. “Do ‘Skills Beget Skills’? Evidence on the Effect of Kindergarten Entrance Age on the Evolution of Cognitive and Non-Cognitive Skill Gaps in Childhood.” *Economics of Education Review* 53: 194–206.
Lucas, Patricia J., Karen McIntosh, Mark Petticrew, Helen M. Roberts, and Alan Shiell. 2008. “Financial Benefits for Child Health and Well-Being in Low Income or Socially Disadvantaged Families in Developed World Countries.” *Campbell Systematic Reviews* 4 (1): 1–93.
Lundberg, Shelly, and Robert A. Pollak. 1993. “Separate Spheres Bargaining and the Marriage Market.” *Journal of Political Economy* 101 (6): 988–1010.
———. 1996. “Bargaining and Distribution in Marriage.” *The Journal of Economic Perspectives* 10 (4): 139–58.
Lundberg, Shelly, Robert A. Pollak, and Jenna Stearns. 2016. “Family Inequality: Diverging Patterns in Marriage, Cohabitation, and Childbearing.” *Journal of Economic Perspectives* 30 (2): 79–102.
Lundborg, Petter, Erik Plug, and Astrid Würtz Rasmussen. 2017. “Can Women Have Children and a Career? IV Evidence from IVF Treatments.” *American Economic Review* 107 (6): 1611–37.
MacKinnon, James G., Morten Ørregaard Nielsen, and Matthew D. Webb. 2023. “Cluster-Robust Inference: A Guide to Empirical Practice.” *Journal of Econometrics* 232 (2): 272–99.
Maclean, Johanna Catherine. 2013. “The Health Effects of Leaving School in a Bad Economy.” *Journal of Health Economics* 32 (5): 951–64.
Mahdi, Samira. 2021. “Innovation Sociale: Évaluation Des Effets d’un Dispositif Innovant Sur Les Professionnels de l’accompagnement Socioprofessionnel: Le Cas Du Projet Reliance Sur Le Département de Meurthe Et-Moselle.” Mémoire de Master Sciences de l’éducation « Formation, travail, Territoires et développement ». Nancy: Université de Lorraine.
Mangin, Evelyn Jean. 2021. “A Narrative Policy Framework and Feminist Critical Policy Analysis of the Welfare

Reform Acts, and The Higher Education of Single Welfare Mothers." PhD thesis, Walden University.

Mani, A., S. Mullainathan, E. Shafir, and J. Zhao. 2013. "Poverty Impedes Cognitive Function." *Science* 341 (6149): 976–80.

Mani, Anandi, Sendhil Mullainathan, Eldar Shafir, and Jiaying Zhao. 2013. "Poverty Impedes Cognitive Function." *Science* 341 (6149): 976–80.

Maniquet, François, and Dirk Neumann. 2021. "Well-Being, Poverty, and Labor Income Taxation: Theory and Application to Europe and the United States." *American Economic Journal: Microeconomics* 13 (2): 276–310.

Manning, Matthew, Susanne Garvis, Christopher Fleming, and Gabriel T. W. Wong. 2017. "The Relationship Between Teacher Qualification and the Quality of the Early Childhood Education and Care Environment." *Campbell Systematic Reviews* 13 (1): 1–82.

Mare, Robert D. 1991. "Five Decades of Educational Assortative Mating." *American Sociological Review* 56 (1): 15–32.

Markussen, Simen, and Knut Røed. 2016. "Leaving Poverty Behind? The Effects of Generous Income Support Paired with Activation." *American Economic Journal: Economic Policy* 8 (1): 180–211.

Marshall, Nancy L. 2004. "The Quality of Early Child Care and Children's Development." *Current Directions in Psychological Science* 13 (4): 165–68.

Martignani, Luca. 2016. "New Risks and the Reframing of Local Welfare as Social Investment: The Case of the FORJAD Program in Switzerland." *Open Journal of Political Science* 06 (01): 10–18.

Martin, Henri, and Hélène Périvier. 2018. "Les Échelles d'équivalence à l'épreuve Des Nouvelles Configurations Familiales:" *Revue Économique* Vol. 69 (2): 303–34.

Martin, John P. 2015. "Activation and Active Labour Market Policies in OECD Countries: Stylised Facts and Evidence on Their Effectiveness." *IZA Journal of Labor Policy* 4 (1): 4.

Martínez, Pablo, José Ignacio Nazif-Munoz, Graciela Rojas, and Irene Magaña. 2022. "Structural Gender Inequalities and Symptoms of Postpartum Depression in 40 Countries." *Journal of Affective Disorders* 297 (January): 381–85.

Masters, N. Tatiana, Taryn P. Lindhorst, and Marcia K. Meyers. 2014. "Jezebel at the Welfare Office: How Racialized Stereotypes of Poor Women's Reproductive Decisions and Relationships Shape Policy Implementation." *Journal of Poverty* 18 (2): 109–29.

Maurin, Eric, and Julie Moschion. 2009. "The Social Multiplier and Labor Market Participation of Mothers." *American Economic Journal: Applied Economics* 1 (1): 251–72.

McCrary, Justin. 2008. "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test." *Journal of Econometrics* 142 (2).

McCRATE, ELAINE, and JOAN SMITH. 1998. "WHEN WORK DOESN'T WORK: The Failure of Current Welfare Reform." *Gender & Society* 12 (1): 61–80.

McGuinness, Seamus. 2006. "Overeducation in the Labour Market." *Journal of Economic Surveys* 20 (3): 387–418.

McLanahan, Sara. 2004a. "Diverging Destinies: How Children Are Faring Under the Second Demographic Transition." *Demography* 41 (4): 607–27.

———. 2004b. "Diverging Destinies: How Children Are Faring Under the Second Demographic Transition." *Demography* 41 (4): 607–27.

Medoff, Marshall H. 2003. "The Impact of Anti-Abortion Activities on State Abortion Rates." *The Journal of Socio-Economics* 32 (3): 265–82.

Meisel, Joan, Daniel Chandler, and Beth Menees Rienzi. 2003. "Domestic Violence Prevalence and Effects on Employment in Two California TANF Populations." *Violence Against Women* 9 (10): 1191–1212.

Melly, Blaise, and Kaspar Wüthrich. 2017. "Local Quantile Treatment Effects." In *Handbook of Quantile Regression*, edited by Roger Koenker, Victor Chernozhukov, Xuming He, and Limin Peng, 1st ed., 145–64. Chapman and Hall/CRC.

Meurs, Dominique, and Pierre Pora. 2019. "Gender Equality on the Labour Market in France: A Slow Convergence Hampered by Motherhood." *Economie Et Statistique* 510 (1): 109–30.

———. 2020. "Gender Equality on the Labour Market in France: A Slow Convergence Hampered by Motherhood." *Economie Et Statistique / Economics and Statistics*, January, 109–30.

Meyer, Bruce D., and Dan T. Rosenbaum. 2001. "Welfare, the Earned Income Tax Credit, and the Labor Supply of Single Mothers*." *The Quarterly Journal of Economics* 116 (3): 1063–1114.

Midgley, James. 1999. "Growth, Redistribution, and Welfare: Toward Social Investment." *Social Service Review*

73 (1): 3–21.
Millar, Michelle, Jane Gray, Clíona Rooney, and Rosemary Crosse. 2018. “The (in)adequacy of in-Work Benefits in Irish Lone Parent Labour Market Activation Policy.” *Journal of Poverty and Social Justice* 26 (3): 379–400.
Milner, Susan. 2017. “Employment and Labour Market Policy During the Hollande Presidency: A Tragedy in Three Acts?” *Modern & Contemporary France* 25 (4): 429–43.
Misra, Joya, Stephanie Moller, and Michelle J. Budig. 2007. “Work—Family Policies and Poverty for Partnered and Single Women in Europe and North America.” *Gender & Society* 21 (6): 804–27.
Moffitt, Robert. 1983. “An Economic Model of Welfare Stigma.” *The American Economic Review* 73 (5): 1023–35.
———. 1992. “Incentive Effects of the U.S. Welfare System: A Review.” *Journal of Economic Literature* 30 (1): 1–61.
Moffitt, Robert A, ed. 1998. *Welfare, the Family and Reprodictive Behavior: Research Perspetives*. Washington DC: NATIONAL ACADEMY PRESS.
Moffitt, Robert A. 2003. “The Negative Income Tax and the Evolution of U.S. Welfare Policy.” *Journal of Economic Perspectives* 17 (3): 119–40.
———. 2015. “The Deserving Poor, the Family, and the U.S. Welfare System.” *Demography* 52 (3): 729–49.
Mogstad, Magne, and Chiara Pronzato. 2012. “Are Lone Mothers Responsive to Policy Changes? Evidence from a Workfare Reform in a Generous Welfare State.” *The Scandinavian Journal of Economics* 114 (4): 1129–59.
Mogstad, Magne, Alexander Torgovitsky, and Christopher R. Walters. 2021. “The Causal Interpretation of Two-Stage Least Squares with Multiple Instrumental Variables.” *American Economic Review* 111 (11): 3663–98.
Molander, Anders, and Gaute Torsvik. 2015. “Getting People into Work: What (If Anything) Can Justify Mandatory Activation of Welfare Recipients?: Getting People into Work.” *Journal of Applied Philosophy* 32 (4): 373–92.
Moller, Stephanie. 2002. “Supporting Poor Single Mothers: Gender and Race in the U.S. Welfare State.” *Gender and Society* 16 (4): 465–84.
Mora, Ricardo, and Javier Ruiz-Castillo. 2011. “Entropy-Based Segregation Indices.” *Sociological Methodology* 41 (1): 159–94.
Morel, Nathalie, Bruno Palier, and Joakim Palme. 2012a. *Towards a Social Investment Welfare State?* Bristol University Press.
———. 2012b. *Towards a Social Investment Welfare State? : Ideas, Policies and Challenges*. Edited by Nathalie Morel, Bruno Palier, and Joakim Palme. Policy Press.
Morescalchi, Andrea, and Paolo Paruolo. 2020. “Too Much Stick for the Carrot? Job Search Requirements and Search Behaviour of Unemployment Benefit Claimants.” *The B.E. Journal of Economic Analysis & Policy* 20 (1).
Morrissey, Taryn W. 2017. “Child Care and Parent Labor Force Participation: A Review of the Research Literature.” *Review of Economics of the Household* 15 (1): 1–24.
Moynihan, Donald, Pamela Herd, and Hope Harvey. 2015. “Administrative Burden: Learning, Psychological, and Compliance Costs in Citizen-State Interactions.” *Journal of Public Administration Research and Theory: J-PART* 25 (1): 43–69.
Myers, Caitlin Knowles. 2017. “The Power of Abortion Policy: Reexamining the Effects of Young Women’s Access to Reproductive Control.” *Journal of Political Economy*, December.
Nafilyan, Vahé. 2016. “Lost and Found?: The Cost of Job Loss in France.” OECD Social, Employment and Migration Working Papers 194.
Negi, Akanksha, and Jeffrey M. Wooldridge. 2021. “Revisiting Regression Adjustment in Experiments with Heterogeneous Treatment Effects.” *Econometric Reviews* 40 (5): 504–34.
Neimanns, Erik. 2022. “Preferences, Vote Choice, and the Politics of Social Investment: Addressing the Puzzle of Unequal Benefits of Childcare Provision.” *Journal of Social Policy* 51 (4): 945–64.
Neyman, Jerzy. 1934. “On the Two Different Aspects of the Representative Method: The Method of Stratified Sampling and the Method of Purposive Selection.” *Journal of the Royal Statistical Society* 97 (4): 558–625.
Nicole-Drancourt, Chantal. 2015. “Mettre En Perspective La Perspective d’investissement Social.” *La Revue de l’Ires* n° 85-86 (2): 171–209.
Nieuwenhuis, Rense. 2020. “The Situation of Single Parents in the EU.” European parliament.
———. 2022. “Single Parents Competing in a Dual-Earner Society: Social Policy to Level the Playing Field.” *The ANNALS of the American Academy of Political and Social Science* 702 (1): 114–28.
Nieuwenhuis, Rense, and Laurie C Maldonado, eds. 2018. *The Triple Bind of Single-Parent Families*. Bristol UK:

Policy Press.
Nollenberger, Natalia, and Núria Rodríguez-Planas. 2015. “Full-Time Universal Childcare in a Context of Low Maternal Employment: Quasi-experimental Evidence from Spain.” *Labour Economics* 36 (October): 124–36.
Nores, Milagros, and W. Steven Barnett. 2010. “Benefits of Early Childhood Interventions Across the World: (Under) Investing in the Very Young.” *Economics of Education Review* 29: 271–82.
Nyman, Pär, Linuz Aggeborn, and Rafael Ahlskog. 2023. “Filling in the Blanks: How Does Information about the Swedish EITC Affect Labour Supply?” *Labour Economics* 85 (December): 102435.
OECD. 2019a. “Educational Attainment by Gender.” OECD Family database.
———. 2019b. “OECD Family Database.” Database. Social Policy Division -Directorate of Employment, Labour and Social Affairs.
Olivetti, Claudia, and Barbara Petrongolo. 2017. “The Economic Consequences of Family Policies: Lessons from a Century of Legislation in High-Income Countries.” *Journal of Economic Perspectives* 31 (1): 205–30.
Oorschot, Wim van. 1991. “Non-Take-Up of Social Security Benefits in Europe.” *Journal of European Social Policy* 1 (1): 15–30.
Oreopoulos, Philip, Till Von Wachter, and Andrew Heisz. 2012. “The Short- and Long-Term Career Effects of Graduating in a Recession.” *American Economic Journal: Applied Economics* 4 (1): 1–29.
Pailhé, Ariane, Lidia Panico, and Marieke Heers. 2020. “Being Born to a Single Mother in France: Trajectories of Father’s Involvement over the First Year of Life.” *Longitudinal and Life Course Studies* 11 (1): 123–49.
Pais, Joana, Agnes Pinter, and Robert F Veszteg. 2011. “COLLEGE ADMISSIONS AND THE ROLE OF INFORMATION: AN EXPERIMENTAL STUDY.” *International Economic Review* 52 (3): 713–37.
Palier, Bruno. 2014. “La Stratégie d’investissement Social,” 75.
Pathak, Parag A. 2011. “The Mechanism Design Approach to Student Assignment.” *Annual Review of Economics* 3 (1): 513–36.
Pathak, Parag A., and Tayfun Sonmez. 2008. “Leveling the Playing Field: Sincere and Sophisticated Players in the Boston Mechanism.” *American Economic Review* 98 (4): 1636–52.
Pathak, Parag A., and Tayfun Sönmez. 2008. “Leveling the Playing Field: Sincere and Sophisticated Players in the Boston Mechanism.” *American Economic Review* 98 (4): 1636–52.
Pavolini, Emmanuele, and Wim Van Lancker. 2018. “The Matthew Effect in Childcare Use: A Matter of Policies or Preferences?” *Journal of European Public Policy* 25 (6): 878–93.
Pavoni, Nicola, and G. L. Violante. 2007. “Optimal Welfare-to-Work Programs.” *Review of Economic Studies* 74 (1): 283–318.
Pearl, Judea. 2009. *Causality Models, Reasoning and Inference*. Cambridge, MA: Cambridge University Press.
Peden, George. 2017. “Liberal Economists and the British Welfare State.” In *Liberalism and the Welfare State: Economists and Arguments for the Welfare State*, Roger E. Backhouse, Bradley W. Bateman, Tamotsu Nishizawa and Dieter Plehwe, 1:39–56. Oxford University Press.
Pega, Frank, Kristie Carter, Tony Blakely, and Patricia J Lucas. 2013. “In-Work Tax Credits for Families and Their Impact on Health Status in Adults.” Edited by Cochrane Public Health Group. *Cochrane Database of Systematic Reviews*, August.
Perez, Sara, Dawn M. Johnson, and Caroline Vaile Wright. 2012. “The Attenuating Effect of Empowerment on IPV-Related PTSD Symptoms in Battered Women Living in Domestic Violence Shelters.” *Violence Against Women* 18 (1): 102–17.
Périvier, Hélène. 2012. “Travaillez Ou Mariez-Vous !la Régulation Sexuée de La Pauvreté En France Et Aux États-Unis.” *Travail, Genre Et Sociétés* 28 (2): 45–62.
———. 2022a. “Étude Sur La Situation Économique Et Sociale Des Parents Isolés : Niveau de Vie, Marché Du Travail Et Politiques Publiques.” Sciences Po.
———. 2022b. *L’économie Féministe*. Les presses Sciences Po.
Périvier, Hélène, and Muriel Pucci. 2019. “Le Recouvrement Des Impayés de Pensions Alimentaires Réduit Les Dépenses Sociales Mais Réduit Également Le Niveau de Vie de Certaines Mères Isolées.” *Sciences Po Publications*, no. info:hdl:2441/2d6dp3s2nh8dmpd40eh37dd6pe (May).
Périvier, Hélène, and Réjane Sénac. 2017. “The New Spirit of Neoliberalism: Equality and Economic Prosperity.” *International Social Science Journal* 67 (223-224): 31–42.
Perrin-Heredia, Ana. 2009. “Les Logiques Sociales de l’endettement : Gestion Des Comptes Domestiques En Milieux Populaires.” *Sociétés Contemporaines* 76 (4): 95–119.
Persky, Joseph. 1997. “Retrospectives: Classical Family Values: Ending the Poor Laws as They Knew Them.”

Journal of Economic Perspectives 11 (1): 179–89.
Peterson, Janice. 1997. “”Ending Welfare as We Know It”: The Symbolic Importance of Welfare Policy in America.” *Journal of Economic Issues* 31 (2): 425–31.
Petitclerc, Amélie, Sylvana Côté, Orla Doyle, Margaret Burchinal, Catherine Herba, Henrik Daae Zachrisson, Michel Boivin, et al. 2017. “Who Uses Early Childhood Education and Care Services? Comparing Socioeconomic Selection Across Five Western Policy Contexts.” *International Journal of Child Care and Education Policy* 11 (1): 3.
Piketty, Thomas, and Emmanuel Saez. 2013. “Optimal Labor Income Taxation.” In *Handbook of Public Economics*, 5:391–474. Elsevier.
Pilet, Jean-Benoit, Lior Sheffer, Luzia Helfer, Frederic Varone, Rens Vliegenthart, and Stefaan Walgrave. 2023. “Do Politicians Outside the United States Also Think Voters Are More Conservative Than They Really Are?” *American Political Science Review*, June, 1–9.
Pirus, Claudine. 2021. “Opinion Des Familles Monoparentales Sur Les Politiques Sociales : Un Sentiment de Vulnérabilité Et Une Attente Forte de Soutien à La Parentalité.” *Études Et Résultats*, DREES, no. 1190.
Pitollat, Claire, and Matthieu Klein. 2018, Aout. “L’accompagnement Des Bénéficiaires Du RSA.” Rapport au 1er Ministre. Paris.
Pollmann-Schult, Matthias. 2005. “Crowding-Out of Unskilled Workers in the Business Cycle: Evidence from West Germany.” *European Sociological Review* 21 (5): 467–80.
———. 2018. “Single Motherhood and Life Satisfaction in Comparative Perspective: Do Institutional and Cultural Contexts Explain the Life Satisfaction Penalty for Single Mothers?” *Journal of Family Issues* 39 (7): 2061–84.
Pora, Pierre. 2020. “Keep Working and Spend Less? Collective Childcare and Parental Earnings in France.” Working Paper. INSEE.
Portela, Mickaël (DREES/SEEE), Céline Marc, Cyrine Hannafi, Rémi Le Gall, Antoine Rode, and Stéphanie Laguérodie. 2022. “Non-Take-up of Minimum Social Benefits: Quantification in Europe.” 94. Paris: DREES.
Prickett, Kate C., and Jennifer March Augustine. 2021. “Trends in Mothers’ Parenting Time by Education and Work From 2003 to 2017.” *Demography* 58 (3): 1065–91.
Pucci, Muriel, and Hélène Périvier. 2022. “Le Soutien Apporté Par Le Système Sociofiscal Aux Parents Séparés : Bilan Et Proposition de Réformes.” *Informations Sociales* 207 (3): 74–82.
Quadrature du net. 2023. “Scoring of Welfare Beneficiaries: The Indecency of CAF’s Algorithm Now Undeniable,” November 2023.
Rabe, Birgitta. 2019. “Do School Inputs Crowd Out Parents’ Investments in Their Children?” *IZA World of Labor*, May.
Redman, Jamie. 2020. “The Benefit Sanction: A Correctional Device or a Weapon of Disgust?” *Sociological Research Online* 25 (1): 84–100.
Reese, Ellen. 2005. *Backlash Against Welfare Mothers: Past and Present*. 1st ed. University of California Press.
Rees-Jones, Alex, and Dmitry Taubinsky. 2020. “Measuring ‘Schmeduling’.” *The Review of Economic Studies* 87 (5): 2399–2438.
Reich-Stiebert, Natalia, Laura Froehlich, and Jan-Bennet Voltmer. 2023. “Gendered Mental Labor: A Systematic Literature Review on the Cognitive Dimension of Unpaid Work Within the Household and Childcare.” *Sex Roles* 88 (11): 475–94.
Reingold, Beth, and Adrienne R. Smith. 2012. “Welfare Policymaking and Intersections of Race, Ethnicity, and Gender in U.S. State Legislatures.” *American Journal of Political Science* 56 (1): 131–47.
Reischmann, Tobias, Thilo Klein, and Sven Giegerich. 2021. “An Iterative Deferred Acceptance Mechanism for Decentralized, Fast and Fair Childcare Assignment.” SSRN Scholarly Paper. Rochester, NY. 2021.
Riger, Stephanie, and Maryann Krieglstein. 2000. “The Impact of Welfare Reform on Men’s Violence Against Women.” *American Journal of Community Psychology* 28 (5): 631–47.
Rinaldi, Chiara, and Marleen P. M. Bekker. 2021. “A Scoping Review of Populist Radical Right Parties’ Influence on Welfare Policy and Its Implications for Population Health in Europe.” *International Journal of Health Policy and Management* 10 (3): 141–51.
Robins, Philip K. 1985. “A Comparison of the Labor Supply Findings from the Four Negative Income Tax Experiments.” *The Journal of Human Resources* 20 (4): 567–82.
Rodrigues, Bruno, and Vincent Vergnat. 2019. “The Time and the Transitions Back to Work in France After Maternity.” *Review of Economics of the Household* 17 (3): 861–88.
Rodríguez, Jorge. 2023. “Understanding the Effects of Workfare Policies on Child Human Capital.” *Journal of*

Labor Economics 41 (1): 39–75.
Rodrik, Dani, and Stefanie Stantcheva. 2021. "Fixing Capitalism's Good Jobs Problem." *Oxford Review of Economic Policy* 37 (4): 824–37.
Rosenbaum, Paul R., and Donald B. Rubin. 1983. "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika* 70 (1): 41–55.
Roth, Alvin E. 1982. "The Economics of Matching: Stability and Incentives." *Mathematics of Operations Research* 7 (4): 617–28.
———. 1984. "The Evolution of the Labor Market for Medical Interns and Residents: A Case Study in Game Theory." *Journal of Political Economy* 92 (6): 991–1016.
———. 2002. "The Economist as Engineer: Game Theory, Experimentation, and Computation as Tools for Design Economics." *Econometrica* 70 (4): 1341–78.
———. 2007. "Repugnance as a Constraint on Markets." *Journal of Economic Perspectives* 21 (3): 37–58.
———. 2015. *Who Gets What–and Why: The New Economics of Matchmaking and Market Design*. Houghton Mifflin Harcourt.
———. 2018. "Marketplaces, Markets, and Market Design." *The American Economic Review* 108 (7): 1609–58.
Roth, Jonathan, Pedro H C Sant'Anna, Alyssa Bilinski, and John Poe. 2021. "What's Trending in Di□erence-in-Di□erences? A Synthesis of the Recent Econometrics Literature," 54.
Roth, Jonathan, Pedro H. C. Sant'Anna, Alyssa Bilinski, and John Poe. 2023. "What's Trending in Difference-in-Differences? A Synthesis of the Recent Econometrics Literature." *Journal of Econometrics* 235 (2): 2218–44.
Rothstein, Jesse, and Till von Wachter. 2017, Mars. "Social Experiment in the Labor Market." In *Handbook of Field Experiments*, 1st ed.
Rubin, Donald B. 1974. "Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies." *Journal of Educational Psychology* 66(5): 688–701.
Saez, Emmanuel. 2010. "Do Taxpayers Bunch at Kink Points?" *American Economic Journal: Economic Policy* 2 (3): 180–212.
Saez, Emmanuel, Joel Slemrod, and Seth H. Giertz. 2012. "The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review." *Journal of Economic Literature* 50 (1): 3–50.
Saez, Emmanuel, and Stefanie Stantcheva. 2016. "Generalized Social Marginal Welfare Weights for Optimal Tax Theory." *American Economic Review* 106 (1): 24–45.
Sant'Anna, Pedro H. C., and Jun Zhao. 2020. "Doubly Robust Difference-in-Differences Estimators." *Journal of Econometrics* 219 (1): 101–22.
Satz, Debra. 2010a. "What Do Markets Do?" In *Why Some Things Should Not Be for Sale: The Moral Limits of Markets*, edited by Debra Satz, 0. Oxford University Press.
———. 2010b. *Why Some Things Should Not Be for Sale: The Moral Limits of Markets*. Edited by Debra Satz. Oxford University Press.
Schilbach, Frank, Heather Schofield, and Sendhil Mullainathan. 2016a. "The Psychological Lives of the Poor." *American Economic Review* 106 (5): 435–40.
———. 2016b. "The Psychological Lives of the Poor." *American Economic Review* 106 (5): 435–40.
Schmauder, Christian, Jurgis Karpus, Maximilian Moll, Bahador Bahrami, and Ophelia Deroy. 2023. "Algorithmic Nudging: The Need for an Interdisciplinary Oversight." *Topoi* 42 (3): 799–807.
Schmidt, Eva-Maria, Fabienne Décieux, Ulrike Zartler, and Christine Schnor. 2023. "What Makes a Good Mother? Two Decades of Research Reflecting Social Norms of Motherhood." *Journal of Family Theory & Review* 15 (1): 57–77.
Schmieder, Johannes F, Till Von Wachter, and Jo Rg Heining. 2022. "The Costs of Job Displacement over the Business Cycle and Its Sources: Evidence from Germany." *THE AMERICAN ECONOMIC REVIEW*.
Schmutz, Rita. 2024. "Is Universal Early Childhood Education and Care an Equalizer? A Systematic Review and Meta-Analysis of Evidence." *Research in Social Stratification and Mobility* 89 (February): 100859.
Schüller, Simone, and Hannah S. Steinberg. 2022. "Parents Under Stress: Evaluating Emergency Childcare Policies During the First COVID-19 Lockdown in Germany." *Labour Economics* 78 (October): 102217.
Sethi, Rajiv, and Rohini Somanathan. 2023. "Meritocracy and Representation." *Journal of Economic Literature* 61 (3): 941–57.
Shah, Anuj K., Jiaying Zhao, Sendhil Mullainathan, and Eldar Shafir. 2018. "Money in the Mental Lives of the Poor." *Social Cognition* 36 (1): 4–19.
Shahidi, Faraz V, Chantel Ramraj, Odmaa Sod-Erdene, Vincent Hildebrand, and Arjumand Siddiqi. 2019. "The

Impact of Social Assistance Programs on Population Health: A Systematic Review of Research in High-Income Countries." *BMC Public Health* 19 (1): 2.
Shutes, Isabel. 2022. "Immigration Policies and the Risks of Single Parenthood for Migrant Women." *The ANNALS of the American Academy of Political and Social Science* 702 (1): 149–62.
Sicsic, Michaël. 2019. "Financial Incentives to Work in France Between 1998 and 2014." *Economie Et Statistique / Economics and Statistics*, January, 13–35.
Simintzi, Elena, Sheng-Jun Xu, and Ting Xu. 2022. "The Effect of Childcare Access on Women's Careers and Firm Performance." SSRN Scholarly Paper. Rochester, NY. November 20, 2022.
Simonnet, Véronique, and Elisabeth Danzin. 2014. "L'effet Du RSA Sur Le Taux de Retour à l'emploi Des Allocataires. Une Analyse En Double Différence Selon Le Nombre Et l'âge Des Enfants." *Economie Et Statistique* 467 (1): 91–116.
Sirugue, Christophe. 2013. "Réforme Des Dispositifs de Soutien Aux Revenus d'activité Modestes." Rapport au Premier Ministre. Paris: Assemblée nationale.
Słoczyński, Tymon. 2022. "Interpreting OLS Estimands When Treatment Effects Are Heterogeneous: Smaller Groups Get Larger Weights." *The Review of Economics and Statistics* 104 (3): 501–9.
Słoczyński, Tymon, S. Derya Uysal, and Jeffrey M. Wooldridge. 2022a. "Abadie's Kappa and Weighting Estimators of the Local Average Treatment Effect." April 15, 2022.
———. 2022b. "Doubly Robust Estimation of Local Average Treatment Effects Using Inverse Probability Weighted Regression Adjustment." August 2, 2022.
Smedslund, G. 2006. "Work Programmes for Welfare Recipients." *Campbell Systematic Reviews*, August.
Solera, Cristina, and Letizia Mencarini. 2018. "The Gender Division of Housework After the First Child: A Comparison Among Bulgaria, France and the Netherlands." *Community, Work & Family* 21 (5): 519–40.
Spencer, Rachael A., Melvin D. Livingston, Briana Woods-Jaeger, Shelby T. Rentmeester, Nolan Sroczynski, and Kelli A. Komro. 2020. "The Impact of Temporary Assistance for Needy Families, Minimum Wage, and Earned Income Tax Credit on Women's Well-Being and Intimate Partner Violence Victimization." *Social Science & Medicine (1982)* 266 (December): 113355.
Stancanelli, Elena G. F. 2008. "Evaluating the Impact of the French Tax Credit on the Employment Rate of Women." *Journal of Public Economics* 92 (10-11): 2036–47.
Stantcheva, Stefanie. 2021. "Why Do We Not Support More Redistribution? New Explanations from Economics Research." In *Combating Inequality*, edited by Olivier Blanchard and Dani Rodrik, 263–70. The MIT Press.
Stavrova, Olga, and Detlef Fetchenhauer. 2015. "Single Parents, Unhappy Parents? Parenthood, Partnership, and the Cultural Normative Context." *Journal of Cross-Cultural Psychology* 46 (1): 134–49.
Steinberg, Hannah Sinja, and Corinna Kleinert. 2022. "Timing of Early Childcare Take-up in Germany: An Application of Rational Choice Theory." *Journal of Marriage and Family* 84 (3): 840–59.
Stevenson, Betsey, and Justin Wolfers. 2006. "Bargaining in the Shadow of the Law: Divorce Laws and Family Distress." *QUARTERLY JOURNAL OF ECONOMICS* 121 (1): 267–88.
———. 2007. "Marriage and Divorce: Changes and Their Driving Forces." *Journal of Economic Perspectives* 21 (2): 27–52.
———. 2009. "The Paradox of Declining Female Happiness." Working Paper. Working Paper Series. May 2009.
Sullivan, Daniel, and Till von Wachter. 2009. "Job Displacement and Mortality: An Analysis Using Administrative Data *." *Quarterly Journal of Economics* 124 (3): 1265–1306.
Sun, Liyang, and Sarah Abraham. 2020. "Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects." *Journal of Econometrics*, December, S030440762030378X.
Tan, Zhiqiang. 2010. "Bounded, Efficient and Doubly Robust Estimation with Inverse Weighting." *Biometrika* 97 (3): 661–82.
Tekin, Erdal. 2007. "Childcare Subsidies, Wages, and Employment of Single Mothers." *Journal of Human Resources* XLII (2): 453–87.
Thomson, Emily. 2009. "Do Ends Justify Means? Feminist Economics Perspectives on the Business Case for Gender Equality in the UK Labour Market." *E-Cadernos CES*, no. 05, 05 (September).
Toossi, Mitra. 2002. "A Century of Change: The U.S. Labor Force, 1950-2050." *Monthly Labor Review*, May, 15–28.
Trump, Kris-Stella. 2020. "When and Why Is Economic Inequality Seen as Fair." *Current Opinion in Behavioral Sciences* 34 (August): 46–51.
Tübbicke, S. 2023. "When to Use Matching and Weighting or Regression in Instrumental Variable Estimation?

Evidence from College Proximity and Returns to College." *Empirical Economics* 65 (6): 2979–99.
Tur-Prats, Ana. 2019. "Family Types and Intimate Partner Violence: A Historical Perspective." *The Review of Economics and Statistics* 101 (5): 878–91.
Tversky, Amos, and Daniel Kahneman. 1974. "Judgment Under Uncertainty: Heuristics and Biases" 185.
Ünver, Özgün, Tuba Bircan, and Ides Nicaise. 2018. "Perceived Accessibility of Childcare in Europe: A Cross-Country Multilevel Study." *International Journal of Child Care and Education Policy* 12 (1): 5.
Vail, Mark I. 2008. "From 'Welfare Without Work' to 'Buttressed Liberalization': The Shifting Dynamics of Labor Market Adjustment in France and Germany." *European Journal of Political Research* 47: 334–58.
Valant, Jon, and Lindsay H. Weixler. 2022. "Informing School-Choosing Families About Their Options: A Field Experiment From New Orleans." *Educational Evaluation and Policy Analysis* 44 (4): 608–37.
Van Doorn, Bas, and Angela Bos. 2019. "Are Visual Depictions of Poverty in the US Gendered and Racialized?" In *The Social Legitimacy of Targeted Welfare: Attitudes to Welfare Deservingness*, 113–26. Edward Elgar.
Van Lancker, Wim. 2013. "Putting the Child-Centred Investment Strategy to the Test: Evidence for the EU27." *European Journal of Social Security* 15 (1): 4–27.
Van Lancker, Wim, and Joris Ghysels. 2016. "Explaining Patterns of Inequality in Childcare Service Use Across 31 Developed Economies: A Welfare State Perspective." *International Journal of Comparative Sociology* 57 (5): 310–37.
Van Winkle, Zachary, and Emanuela Struffolino. 2018. "When Working Isn't Enough: Family Demographic Processes and in-Work Poverty Across the Life Course in the United States." *Demographic Research* 39 (September): 365–80.
Vandenbroucke, Frank, and Koen Vleminckx. 2011a. "Disappointing Poverty Trends: Is the Social Investment State to Blame?" *Journal of European Social Policy* 21 (5): 450–71.
———. 2011b. "Disappointing Poverty Trends: Is the Social Investment State to Blame?" *Journal of European Social Policy* 21 (5): 450–71.
Verlaat, Timo, Stephanie Rosenkranz, Loek F. M. Groot, and Mark Sanders. 2021. "Requirements Vs. Autonomy: What Works in Social Assistance?" SSRN Scholarly Paper. Rochester, NY. June 21, 2021.
Vermare, Sébastien, Bérengère Junod-Mesqui, Marion Cochard, and Franck Arnaud. 2008. "Les Effets Incitatifs de La Prime Pour l'emploi : Une Évaluation Difficile." *Economie Et Statistique* 412 (1): 57–80.
Verniers, Catherine, Virginie Bonnot, and Yvette Assilaméhou-Kunz. 2022. "Intensive Mothering and the Perpetuation of Gender Inequality: Evidence from a Mixed Methods Research." *Acta Psychologica* 227 (July): 103614.
Veski, Andre, Peter Biro, Kaire Poder, and Triin Lauri. 2017. "Efficiency and Fair Access in Kindergarten Allocation Policy Design." *Journal of Mechanism and Institution Design* 2 (1): 57–104.
Vooren, Melvin, Carla Haelermans, Wim Groot, and Henriëtte Maassen van den Brink. 2019a. "The Effectiveness of Active Labor Market Policies: A Meta-Analysis." *Journal of Economic Surveys* 33 (1): 125–49.
———. 2019b. "The Effectiveness of Active Labor Market Policies: A Meta-Analysis." *Journal of Economic Surveys* 33 (1): 125–49.
Wachter, Till von. 2020. "The Persistent Effects of Initial Labor Market Conditions for Young Adults and Their Sources." *Journal of Economic Perspectives* 34 (4): 168–94.
Waldinger, Daniel. 2021. "Targeting In-Kind Transfers Through Market Design: A Revealed Preference Analysis of Public Housing Allocation." *American Economic Review* 111 (8): 2660–96.
Walters, Christopher R. 2015. "Inputs in the Production of Early Childhood Human Capital: Evidence from Head Start." *American Economic Journal: Applied Economics* 7 (4): 76–102.
Weixler, Lindsay, Jon Valant, Daphna Bassok, Justin B. Doromal, and Alica Gerry. 2020a. "Helping Parents Navigate the Early Childhood Education Enrollment Process: Experimental Evidence From New Orleans." *Educational Evaluation and Policy Analysis* 42 (3): 307–30.
———. 2020b. "Helping Parents Navigate the Early Childhood Education Enrollment Process: Experimental Evidence From New Orleans." *Educational Evaluation and Policy Analysis* 42 (3): 307–30.
White, Howard. 2019. "The Twenty-First Century Experimenting Society: The Four Waves of the Evidence Revolution." *Palgrave Communications* 5 (1): 47.
Whitmore Schanzenbach, Diane, and Michael R. Strain. 2021. "Employment Effects of the Earned Income Tax Credit: Taking the Long View." *Tax Policy and the Economy* 35 (June): 87–129.
Will, J. A. 1993. "The Dimensions of Poverty: Public Perceptions of the Deserving Poor." *Social Science Research* 22 (3): 312–32.
Windebank, J. 2001. "Dual-Earner Couples in Britain and France: Gender Divisions of Domestic Labour and

Parenting Work in Different Welfare States." *Work, Employment and Society* 15 (June): 269–90.

Wolfers, Justin. 2006. "Did Unilateral Divorce Laws Raise Divorce Rates? A Reconciliation and New Results." *American Economic Review* 96 (5).

Wunsch, Conny. 2016. "How to Minimize Lock-in Effects of Programs for Unemployed Workers." *IZA World of Labor*, January.

Wüthrich, Kaspar, and Ying Zhu. 2023. "Omitted Variable Bias of Lasso-Based Inference Methods: A Finite Sample Analysis." *The Review of Economics and Statistics* 105 (4): 982–97.

Zangger, Christoph, and Janine Widmer. 2020. "Choosing What Is Best for One's Children? Experimental Evidence on Parents' Responsiveness to Childcare Subsidies and Their Preferences for Different Childcare Arrangements." *Early Childhood Research Quarterly* 51 (April): 110–23.

Ziliak, James P. 2015. "Temporary Assistance for Needy Families." In *Economics of Means-Tested Transfer Programs*, edited by Robert A. Moffitt, University of Chicago Press. Chicago.

Zoch, Gundula. 2020. "Public Childcare Provision and Employment Participation of East and West German Mothers with Different Educational Backgrounds." *Journal of European Social Policy* 30 (3): 370–85.