Random redistribution and discrimination*

Stéphane Gauthier†
PSE, University of Paris 1 and Institute for Fiscal Studies
Guy Laroque‡
Sciences-Po, UCL and Institute for Fiscal Studies

June 13, 2024

Abstract

We examine optimal random nonlinear income taxation in a Mirrlees economy with risk-averse agents and quasilinear utility in labor. Bunching in the deterministic optimum is necessary for socially beneficial random redistribution. Fiscal discrimination that is unattainable in the deterministic case due to bunching, becomes feasible in the stochastic case. Circumstances favoring lottery-based redistribution over deterministic policies involve a shift from a downward to an upward pattern of incentives, aligning redistribution goals with incentives. In an example with a Rawlsian government, lotteries should be allocated to the most risk-averse population rather than the least risk-averse.

JEL classification numbers: C61, D82, D86, H21.

Keywords: redistribution, asymmetric information, bunching, random taxation.

---

*We are grateful to Francis Bloch, Ben Brooks, Killian Dengreville, Piotr Dworczak, Marc Fleurbaey, Deniz Kattwinkel, David Martimort, Nicolas Werquin, Floris Zoutman and participants to the Public Finance and Mechanism Design session of the 2023 conference of the Society for the Advancement of Economic Theory in Paris for feedback on our work. The usual disclaimers apply.

†stephane.gauthier@univ-paris1.fr; PSE, 48 bd Jourdan, 75014 Paris, France.

‡g.laroque@ucl.ac.uk; UCL, Drayton House, 30 Gordon St, London WC1H 0AX, UK.

1 Introduction

In a first-best environment the government is assumed to have the ability to observe the innate traits of individuals, which can serve as a basis for the design of redistribution policies. Financial assistance, such as food stamps, housing subsidies or cash benefits can be provided to the low-skilled poor, disabled persons or those with qualifying medical conditions preventing gainful activity. The spectrum of policies available to the government often is more limited when some relevant characteristics of individuals are not publicly known. Extra resources then have to be spent to perform a suitable targeting of income support to those in need, implying a balance between efficiency and equity concerns. An important insight of Mirrlees (1971) is to provide us with a formal representation of the additional costs due to asymmetric information. When relevant traits are privately known to potential recipients, the government must also take into account incentive constraints to meet self-selection among those genuinely in need of support, while deterring claims from those who do not require assistance.

The literature shows that the burden of asymmetric information typically falls on those in need, rather than those in more favorable situations. Indeed the typical response to asymmetric information is to reduce the amount of assistance. The lower aid enables the government to target those in need as beneficiaries of assistance while others are discouraged.

Several ideas have therefore been explored to expand possible redistribution through improved targeting of assistance. Most of them involve some form of ordeal mechanism, in the spirit of Nichols and Zeckhauser (1982). The general idea to achieve a better targeting is to subject vulnerable populations to challenging tasks or stressful situations. The government can for instance rely on time-consuming shameful queuing to distribute essential goods to low-income households. It may implement unnecessarily complex and lengthy application processes to prove eligibility for benefits, or impose additional conditions after admission to continue receiving benefits, such as requiring beneficiaries to regularly send their children to school or undergo health check-ups. The social usefulness of these complementary schemes depends on balancing the direct cost borne by the targeted population and the benefits from relaxed incentive constraints associated with the discouragement of undue claims.

In these examples ordeal is usually taken as deterministic, i.e., pain comes for sure. In this paper, we are interested into a specific form of ordeal, which is to impose random noisy transfers to risk averse recipients. It is known from Weiss (1976), Stiglitz (1981), Stiglitz (1982) or Brito, Hamilton, Slutsky, and Stiglitz (1995) that deterministic redistribution sometimes is socially dominated. In Lang (2017), Ederer, Holden, and Meyer (2018) or Lang (2023), randomization limits gaming social rules by blurring incentives. Vague standards or legal uncertainty deter firms to undertake strategies detrimental to the society. Hospitals, for instance, may be discouraged from selecting healthy but less costly to treat patients if they are not aware of the exact amount of compensation they will receive.

In public finance, income lotteries follow from random noise in taxes, because of e.g., administration errors, tax evasion coupled with non-comprehensive auditing, or uncertainty about the actual fiscal regime amid frequent tax reforms. In Wijkander (1988) or Dworczak, Kominers, and Akbarpour (2021), lotteries occur in the presence of quotas and rationing in the allocation of certain goods or services, as limits on market transactions lead some agents to engage in trade with strictly interior probabilities (with some risk of being rationed). In the same vein, random labor and before-tax income variations can be induced by the minimum wage and the risk of unemployment; they can also be due to randomness in occupations for students who apply for medical training in the Netherlands and are accepted by draw. But perhaps, rather than before-tax income, the most explicit randomizations concern situations where the government instead relies on random after-tax incomes and allocation of consumption goods for redistributive purposes. Tobin (1971) argues that income tests for housing subsidies make support 'available only for an accidentally or arbitrarily selected few' while randomness from housing programs involving rent regulation improves selectivity in access for low-income populations in Weitzman (1977). Similar situations are actually very common in situations where an agency has to allocate scarce resources. For instance, in the distribution of public piped water and energy resources in urban areas of developing countries, authorities frequently handle shortages through rationing, potentially employing random allocation methods.

Solutions to optimization programs that we use in economics to characterize optimal policies can involve randomness because of failures of standard convexity assumptions. In the presence of asymmetric information, these failures come from the fact that the allocation intended for a given agent influences both her utility and the utility of those who are willing to mimic her, thus appearing in both sides of the incentive constraints. Randomizing over allocations located on the frontier of the set delimited by the constraints leads to allocations on the convex hull of the constraints of the deterministic program. The deterministic and random constraint sets then differ, which makes it possible for some allocations within the convex hull to outperform the best deterministic alternative.[1]

The general flavor of the economic argument seems very simple to grasp. Following Hellwig (2007), suppose that the government would like to redistribute income to low-skilled in a population of risk averse workers. Redistribution is potentially limited if the government observes neither skill nor the exact amount of labor, as high-skilled then might reduce labor effort to enjoy higher transfers. Ordeal from randomness in the after-tax income designed for low-skilled is detrimental to their welfare, but this also expands the scope of possible redistribution by discouraging risk-averse high-skilled from relaxing effort. A deterministic optimum obtains if the welfare cost incurred by those facing noise overcomes the gain from expanded scope of redistribution. This is more likely to happen if high-skilled do not suffer much from income noise. Hellwig (2007) indeed shows that the government should rely on deterministic redistribution if risk aversion decreases with labor productivity, i.e., risk aversion is higher among low-skilled individuals than among high-skilled individuals. This result leaves us with little hope for randomized taxes to improve the welfare of the poor, who are usually found more risk averse than the rich.

Although the above argument for useful randomization sounds highly intuitive, it does not accord with a puzzling parametric example in Strausz (2006). In this example, the first-best policy of a regulatory authority violates incentive compatibility, fitting the familiar pattern with high production cost (inefficient) firms ready to mimic low cost (efficient) firms. The profit of inefficient firms displays greater concavity, implying greater risk aversion to random production requirements that would be set by the authority. Nevertheless, the second-best regulatory policy involves a random option designed for the inefficient firms, which are the most rather than the less risk-averse firms.

The counter-example in Strausz (2006) suggests that the economic argument identified so far for the role played by randomization does not fully account for the role of random noise in the presence of asymmetric information. Our paper provides an example in the same vein as Strausz (2006) in

[1] See also Pavlov (2011), Gauthier and Laroque (2014), Pycia and Unver (2015) or Gauthier and Laroque (2017) for related approaches.

an optimal taxation setup with a continuum of risk-averse taxpayers. Our example puts forward a new channel for socially useful random policies. It combines bunching in the deterministic optimum and a reversal of incentives.

We consider a Rawlsian government that only values the agents with the lowest utility. If incentive compatibility issues could be dealt with the first-order approach that neglects the possibility of bunching, the best deterministic redistribution policy would involve socially disfavored (rich) types envying the option designed for those more socially favored (poor). However, the strong redistribution motive underlying Rawlsian criteria leads to bunching where many different taxpayers, including those socially favored, have to enjoy the same income transfers and eventually earn the same after-tax income. Thus, even though government would like to discriminate taxpayers, this is prevented for incentive reasons. Eventually no redistribution relying on deterministic tax tools is possible.

The uniform treatment of the agents in the deterministic optimum with bunching akin to some form of uniform rationing blurs the pattern of incentives. In this situation, every agent affected by bunching may be seen as both unwilling to mimic any other agent and envied by the others. Small tax randomization can then allow the government to exploit heterogeneity in risk aversion in a way that reverses the pattern of incentives compared to the deterministic case, where risk aversion does not matter. We exhibit random transfers making the agents that the government wants to favor now envying the treatment of those with lower social importance, a feature reminiscent to countervailing incentives. That is, tax randomization allows for fiscal discrimination by aligning incentives with the social desire for redistribution. A similar reversal occurs in Strausz (2006), but not in Hellwig (2007) where the same structure of incentives prevails both in the deterministic and stochastic cases. Actually it should be clear that the disappointing outcome for pro-poor policies in Hellwig (2007) relies on the fact that high-skilled types continue to envy the low skilled once random noise is introduced into the tax system. Our example shows that this is not a general property.

The gain from the aligned incentives however comes with a cost, as the agents favored by the government have to face randomness. In a particular parametrization of our model the gain from aligned incentives overcomes the cost, and so redistribution should involve a random income for the socially favored (lowest utility) agents, though they display the highest risk aversion. This may provide incentive-based justifications for the risk of unemployment induced by the minimum wage, and other forms of rationing coupled with random allocation commonly used in social assistance and housing policies.

The paper proceeds as follows. Our setup with random redistribution is described in Section 2. Section 3 characterizes the role played by bunching in the deterministic optimum. Section 4 shows that, in the presence of small random noise in taxes, the direction of incentives can be reversed compared to the deterministic case. Section 5 provides a condition for socially useful randomization in the special case where bunching prevents any deterministic redistribution. Some properties of random redistribution in this case are discussed in Section 6. The analysis is generalized in Section 7, and Section 8 presents a parametric example where optimal redistribution involves randomness. The gain from small randomization appears as quantitatively modest in this example. Finally, Section 9 concludes.

2 General framework

A government wants to redistribute income between a continuum of agents in a population of total unit size. Heterogeneity across agents is characterized by θ, a real parameter taking values in $\Theta = \left[\theta^{\inf}, \theta^{\sup}\right]$, which is referred to as the type of the agent. It has cumulative distribution function $F : \Theta \to [0,1]$ associated with positive density $f : \Theta \to \mathbb{R}_{++}$.

The preferences of a type θ agent are represented by the quasilinear utility function

$$u(c,\theta) - y. \tag{1}$$

The quasilinear formulation is often used in industrial organization and contract theory where c is a quantity of some good sold to a valuation type θ buyer against a total payment of y. One can also view c as benefits in, e.g., public housing and rent assistance voucher programs, while y is the contribution or rent made to the regulatory agency. In an interpretation more in line with the public finance literature, y is before-tax labor income, the government levies the tax $y - c$, and c is after-tax labor income that is also consumption. Earning y requires some effort, hence the disutility cost.

The function u is increasing and differentiable everywhere in c and θ. It is also strictly concave in c, so that every agent is risk-averse. Its first derivative $u'_c(c,\theta)$ with respect to c coincides with the marginal rate of substitution of consumption for before-tax income for a type θ agent. It is assumed to satisfy the Spence-Mirrlees condition that $u'_c(c,\theta)$ is decreasing in θ for all (c,θ), i.e., agents with a higher type value an extra amount of consumption less.

The formulation (1), where θ enters as an argument of the function u, suggests interpreting θ as related to consumption tastes. In view of the direction of the Spence-Mirrlees assumption, it may also be interpreted as accounting for other incomes from, e.g., intra-family transfers or inherited capital, or some extra access to housing facilities. Higher types are better endowed in these other incomes or capacities and thus value an additional amount of good less.

This latter interpretation may be the most natural one in the parametric example developed in Section 8, where $u(c, \theta)$ is specialized to $\ln(c + \theta)$, so that θ must be expressed in units of after-tax income. The additive formulation may also fit the case of housing, with c and θ being the surface areas allocated by the agency and the access capacities available to the recipients, respectively. An agency with redistribution concerns would probably seek to give less housing to those who are otherwise better off, for example, young adults more helped by their parents. An egalitarian objective could go as far as making the total surface area (the sum of the distributed surface area c and θ) the same for everyone. More generally, θ can stand for some substitute to the distributed good c against payment y. Think of grey water, stored rainwater or groundwater that are substitutes for public piped water, or various kinds of solar panels that can be used to substitute for publicly provided electricity.

It is important to emphasize that the direction of our Spence-Mirrlees condition departs from the main stand of the public finance literature, where θ is usually taken as a labor productivity parameter, implying a positive cross-derivative for $u(c, \theta)$. In, e.g., Lollivier and Rochet (2003), an agent with labor productivity θ must provide a labor effort y/θ to earn y. Her utility is $v(c) - y/\theta$ or, after multiplying by θ, $\theta v(c) - y$. So $u(c, \theta) = \theta v(c)$ and $u''_{c\theta}(c, \theta) = v'(c) > 0$. In the economic examples discussed, however, the negative derivative in our Spence-Mirrlees condition seems more appropriate. It is more likely that the marginal utility of additional living space from a housing assistance program is lower for those with greater outside housing access, and public piped water and electricity should be less valuable for those who have better access to private alternatives.

The government designs a redistribution policy between the agents. The policy is defined by a menu $(\tilde{c}(\theta), \tilde{y}(\theta))$ of, say, after- and before-tax income lotteries. The menu is feasible if aggregate consumption falls below aggregate production,

$$\int_{\Theta} \mathbb{E}[\tilde{c}(\theta) - \tilde{y}(\theta)] \, \mathrm{d}F(\theta) \leq 0. \tag{2}$$

The government is assumed to know the distribution of types, but not to observe the value of θ for every agent. This value remains private information to the agent. Therefore the government must also ensure that agents choose the income pair designed for them. This is satisfied if the incentive constraints

$$\mathbb{E}[u(\tilde{c}(\theta), \theta) - \tilde{y}(\theta)] \geq \mathbb{E}[u(\tilde{c}(\tau), \theta) - \tilde{y}(\tau)] \tag{3}$$

hold for all (θ, τ) in $\Theta \times \Theta$.

An optimal redistribution policy is a menu $(\tilde{c}(\theta), \tilde{y}(\theta))$ that maximizes the social welfare objective of the government subject to the feasibility constraint (2) and the incentive constraints (3).

We expect the wealthier consumers to have better access to private substitutes. See Abajian et al. (2024) for recent empirical evidence on water supply in Cape Town. Therefore, it seems natural for a redistributive government to prioritize individuals with low, more likely less well-off types. Here, we consider a Rawlsian government that only values the type of agents with the lowest utility. In all the paper, the type with the lowest utility appears to be $\theta^{\inf}$. Let $V(\theta) = \mathbb{E}[u(\tilde{c}(\theta), \theta) - \tilde{y}(\theta)]$ denote the expected indirect utility of type θ when she chooses the contract designed for her. The Ralwsian social objective is

$$V(\theta^{\inf}). \tag{4}$$

The policy is deterministic if it is made of degenerate lotteries $(c(\theta), y(\theta))$ where every type θ earns the before-tax income $y(\theta)$ with certainty and consumes $c(\theta)$ with certainty (the absence of a tilde mark applies to deterministic options). We are interested into circumstances where some agents face non-degenerate lotteries in the optimal redistribution policy.

3 Randomness and uniform rationing

Given the quasilinear form of agents' utility, replacing the lottery $\tilde{y}(\theta)$ with the sure outcome $y(\theta) = \mathbb{E}[\tilde{y}(\theta)]$ affects neither the constraints nor the social objective. Thus there is no loss to consider that every type θ earns before-tax $y(\theta)$ with certainty. In this section, we show that a menu of after-tax income lotteries $(\tilde{c}(\theta))$ can be socially useful only if incentive considerations prevent the government relying on deterministic fiscal tools from discrimination. That is, a uniform tax treatment has to be applied to different types of agents.

Following the terminology used in, e.g., Laffont and Martimort (2002), we define the optimal 'relaxed' redistribution policy as a menu $(\tilde{c}(\theta), y(\theta))$ maximizing the social objective (4) subject to the feasibility constraint (2) and the necessary first-order conditions for a local truthful report in (3),

$$V'(\theta) = \mathbb{E}\left[u'_\theta(\tilde{c}(\theta), \theta)\right] \tag{5}$$

for all θ. The optimal relaxed redistribution policy coincides with the optimal redistribution policy if it meets the incentive constraints (3), but not otherwise. Summing up the first-order conditions yields

$$V(\theta) = V(\theta^{\inf}) + \int_\Theta \mathbb{E}\left[u'_\theta(\tilde{c}(z), z)\right] \, \mathrm{d}z.$$

Replacing $y(\theta)$ with $\mathbb{E}[u(\tilde{c}(\theta), \theta)] - V(\theta)$ into the feasibility constraint (2) gives the indirect utility $V(\theta^{\inf})$. After using the integration by parts formula, it writes as

$$V(\theta^{\inf}) = \int_\Theta \mathbb{E}[W(\tilde{c}(\theta), \theta)] \, \mathrm{d}F(\theta) \tag{6}$$

where $m(\theta) = [1 - F(\theta)] / f(\theta)$ is the Mills ratio and

$$W(c, \theta) = u(c, \theta) - c - m(\theta) u'_\theta(c, \theta)$$

represents the virtual contribution of type θ to social welfare when she faces after-tax income c with certainty.

In the optimal deterministic relaxed redistribution policy, every type θ gets for sure

$$c^*(\theta) = \arg\max_c W(c, \theta).$$

Suppose that this policy coincides with the optimal deterministic redistribution policy. Then, in the deterministic case, the necessary first-order conditions for a local truthful report are sufficient to deal with incentive requirements. There is no bunching in the deterministic optimum and, in view of the Spence-Mirrlees assumption, $c^*(\theta)$ is non-increasing in θ. The government can discriminate agents by designing a profile of transfers that make the after-tax labor income lower for higher types, those with lower social importance, who put less value on consumption.

Since then $W(c^*(\theta), \theta) \geq W(c, \theta)$ for all c and θ, we have $W(c^*(\theta), \theta) \geq \mathbb{E}[W(\tilde{c}, \theta)]$ for all $\tilde{c}$ and θ. In particular, the inequality holds true if for all θ we set $\tilde{c}$ equal to the lottery $\tilde{c}(\theta)$ that maximizes (6). This yields:

Lemma 1. *A random redistribution policy is socially useless if the optimal deterministic redistribution policy coincides with the optimal deterministic relaxed policy, i.e., the optimal deterministic redistribution policy involves no bunching.*

Proof. The argument given above leads to

$$\int_{\Theta} \mathbb{E}[W(\tilde{c}(\theta), \theta)] \, \mathrm{d}F(\theta) \leq \int_{\Theta} W(c^*(\theta), \theta) \, \mathrm{d}F(\theta)$$

for every menu $(\tilde{c}(\theta))$. To conclude the proof, observe that social welfare in an optimal redistribution policy cannot be greater than welfare in an optimal relaxed policy (which is in the left-hand side of the above inequality). □

Lemma 1 shows that socially useful randomness in redistribution can be achieved only if the incentive constraints associated with the optimal deterministic policy imply bunching, where different types of agents face the same income option. In the context of housing, the authority would be compelled to allocate similar housing goods to recipients with different, but privately known, outside access capacities to housing summarized by θ.

Such an outcome can be interpreted as a form of uniform rationing. Indeed, in a market where all agents would face the same price, those with greater outside access capacities (θ is high) would purchase less housing than others (θ small), as they value less extra housing from the market (the marginal utility $u'_c(c, \theta)$ decreases with θ). The former type of agents may accordingly be viewed as being forced to have greater consumption, while the latter are rationed. This suggests a possible role for randomness as a discriminatory tool for the government. Random transfers are beneficial to the extent that a non-uniform treatment of agents is socially desirable but cannot be achieved using deterministic tools. This is the idea we explore in the next section.

4 A reversal of incentives

If type θ faces the pair $(c(\theta), y(\theta))$ with certainty, the incentive constraints (3) simplify to

$$V(\theta) = u(c(\theta), \theta) - y(\theta) \geq u(c(\tau), \theta) - y(\tau)$$

for all θ and τ. Type θ must receive the informational rent $u(c(\tau), \theta) - u(c(\tau), \tau)$ so that she does not mimic type τ. This is a positive amount for $\tau < \theta$ since $u'_\theta(c, \theta) > 0$ for all (c, θ).

Following textbook arguments in, e.g., Laffont and Martimort (2002), the menu $(c(\theta), y(\theta))$ satisfies the incentive constraints if and only if, for all θ, $V'(\theta) = u'_\theta(c(\theta), \theta)$ and the second-order monotonicity requirement that $c(\theta)$ is non-increasing. Incentive considerations imply an indirect utility increasing with type, $V'(\theta) \geq 0$ for all θ, because of informational rents. The Rawlsian government, if relying on deterministic tax tools, cares about the lowest type $\theta^{\inf}$ of agents only.

Our argument in favor of randomization is easier to grasp in the configuration where the best 'relaxed' menu $(c^*(\theta))$ violates monotonicity requirements for incentive compatibility, and eventually every agent is concerned by bunching in the deterministic optimum (more general configurations are examined in Section 7). In this polar case, every agent receives the same after-tax income c^*. From (6), in the absence of random noise, we have

$$c^* = \arg\max_c \int_\Theta W(c, \theta) \; \mathrm{d}F(\theta).$$

Applying the integration by parts formula,

$$\int_\Theta m(\theta) u'_\theta(c, \theta) \; \mathrm{d}F(\theta) = u(c, \theta^{\inf}) + \int_\Theta u(c, \theta) \; \mathrm{d}F(\theta),$$

so that $V(\theta^{\inf}) = u(c, \theta^{\inf}) - c$ and the optimal level c^* of income, if interior, satisfies

$$u'_c\left(c^*, \theta^{\inf}\right) = 1. \tag{7}$$

The incentive constraints require that agents also earn the same before-tax income y^*. Hence the utility obtained by each type θ agent is $u(c^*, \theta) - y^*$. This alternative can be implemented in a decentralized setup through the use of an income tax schedule with two brackets, where before-tax income is taxed at a very low rate if falling below y^* while it is taxed a high enough rate otherwise. Since feasibility requires $y^* = c^*$, there is no redistribution at all in the optimal deterministic policy. This completes the characterization of this policy.

Let us now randomize the after-tax income. A lottery $\tilde{c}(\theta) = c^* + \tilde{\varepsilon}(\theta)$ is designed for type θ. We consider type-specific realizations of the random variable $\tilde{\varepsilon}(\theta)$ that stand close to 0. In addition, we require that greater noise comes with greater income transfer. The mean and variance of $\tilde{\varepsilon}(\theta)$ are so that $\mathbb{E}[\tilde{\varepsilon}(\theta)] = \text{var}[\tilde{\varepsilon}(\theta)] = \lambda v(\theta)$, with λ a positive real number close to 0 and $v(\theta) \geq 0$ a (rescaled) variance bounded from above. For such lotteries, the (second-order Taylor expansion of the) expected utility of type θ when she chooses the lottery $\tilde{c}(\tau)$ designed for type τ writes

$$\mathbb{E}\left[u\left(c^* + \tilde{\varepsilon}(\tau), \theta\right)\right] \simeq u\left(c^*, \theta\right) + \lambda u_c'\left(c^*, \theta\right)\left(1 - \frac{A\left(c^*, \theta\right)}{2}\right) v(\tau)$$

where

$$A(c, \theta) = -\frac{u_{cc}''(c, \theta)}{u_c'(c, \theta)} > 0$$

is the coefficient of absolute risk aversion of type θ. Randomness in after-tax income allows the government to exploit heterogeneity in individual risk aversion, captured by $A\left(c^*, \theta\right)$. Indeed randomness implies a change in the sub-utility $u\left(c^*, \theta\right)$ obtained from consumption in the deterministic case that is $\lambda S\left(c^*, \theta\right) v(\tau)$ for type θ, with

$$S\left(c^*, \theta\right) = u_c'\left(c^*, \theta\right)\left(1 - \frac{A\left(c^*, \theta\right)}{2}\right).$$

The quantity $S(c, \theta)$ approximates the change in utility derived by type θ agent following a unit increase in her after-tax income from c. Hence it may be thought of as a valuation for the consumption good. In view of (7) and the Spence-Mirrlees assumption, we have

$$S\left(c^*, \theta\right) \leq u_c'(c^*, \theta^{\inf}) + \frac{1}{2} u_{cc}''\left(c^*, \theta\right) < 1.$$

The shape of $S(c, \theta)$ is otherwise difficult to characterize. The interpretation in terms of extra utility suggests that the relevant economic case should be one where $S(c, \theta)$ takes positive values while $u''_{cc}(c, \theta) < 0$. In addition, we also expect from the direction of the Spence-Mirrlees assumption some pressures for the utility gain to be decreasing in θ, $S'_\theta(c^*, \theta) < 0$. This is indeed what happens for standard specifications.

Example 1. CRRA preferences. Suppose that $u(c, \theta) = c^{1-\theta}/(1-\theta)$ for $\theta \neq 1$. The characterization of the optimal after-tax income in (7) gives $c^* = 1$. When evaluated at this income, it is readily checked that

$$S(c^*, \theta) = 1 - \frac{\theta}{2} \text{ and } S'_\theta(c^*, \theta) = -\frac{1}{2} < 0.$$

Referring to the upper bound for the coefficient θ of relative risk aversion obtained by Chetty (2006), which equals 2, we have $S(c^*, \theta) > 0$ for empirically plausible values of θ.

Example 2. CARA preferences. Let $u(c, \theta) = -\exp(-\theta(c - \bar{c}))/\theta$ for some consumption $\bar{c} \geq 0$. Then (7) gives $c^* = \bar{c}$, and both $S(c^*, \theta)$ and $S'_\theta(c^*, \theta)$ are as in the CRRA case. The estimates of the coefficient of absolute risk aversion θ in Cohen and Einav (2007) are of an order of magnitude of 10^{-2} at most. For such values, $S(c^*, \theta) > 0$.

Example 3. Logarithmic Preferences. Let $u(c, \theta) = \ln(c + \theta)$. Then (7) gives $c^* = 1 - \theta^{\inf}$. Non-negative consumption requires $\theta^{\inf} \leq 1$. With $\theta^{\inf} = 0$,

$$S(c^*, \theta) = \frac{1}{c^* + \theta}\left(1 - \frac{1}{2}\frac{1}{c^* + \theta}\right) > 0,$$

and

$$S'_\theta(c^*, \theta) = -\frac{1}{(c^* + \theta)^2}\left(1 - \frac{1}{c^* + \theta}\right) < 0.$$

From now onward, based on the insights from these examples, we assume:

Assumption A1. *The after-tax income valuation $S(c^*, \theta)$ takes positive values and decreases with θ.*

Let $y(\theta)$ denote the before-tax income of type θ in the presence of after-tax income perturbations; it is $y^* + \mathrm{d}y(\theta)$ for some deterministic $\mathrm{d}y(\theta)$ close to 0. The indirect utility $V(\theta)$ writes

$$\mathbb{E}\left[u\left(c^* + \tilde{\varepsilon}(\theta), \theta\right)\right] - y(\theta) \simeq u\left(c^*, \theta\right) + \lambda S\left(c^*, \theta\right) v(\theta) - y(\theta) \tag{8}$$

of type θ under thruthful reporting. The incentive constraints (3) then reduce to

$$U(\theta) = \lambda S\left(c^{*}, \theta\right) v(\theta) - y(\theta) \geq \lambda S\left(c^{*}, \theta\right) v(\tau) - y(\tau)$$

for all τ and θ. Incentive constraints are now driven by a sub-utility $U(\theta)$ of the overall utility $V(\theta)$ of the agents, with monotonicity properties that may differ from those of $V(\theta)$. This is what makes possible a reversal of incentives when random noise is introduced into the redistribution policy.

From a formal viewpoint, our parametrization of lotteries gives rise to incentive constraints which have a familiar shape in $v(\theta)$ and $y(\theta)$. It is then straightforward to deal with incentives in the presence of tax lotteries. The usual arguments yield:

Lemma 2. *Consider a menu where the government uses random perturbations $\tilde{\varepsilon}(\theta)$ to the optimal deterministic after-tax income c^{*} with mean and variance $\lambda v(\theta)$ for some $\lambda \geq 0$ close to 0. The incentive constraints (3) are satisfied if and only if*

$$U'(\theta) = \lambda S'_{\theta}\left(c^{*}, \theta\right) v(\theta)$$

and $v(\theta)$ is non-increasing for all θ.

Proof. Using the envelope theorem, a necessary first-order condition for a local truthful report is $U'(\theta) = \lambda S'_{\theta}\left(c^{*}, \theta\right) v(\theta)$ for all θ. The second-order conditions write $S'_{\theta}\left(c^{*}, \theta\right) v'(\theta) \geq 0$ for all θ where v is differentiable. Finally,

$$\frac{\partial}{\partial \tau}\left(S\left(c^{*}, \theta\right) v(\tau) - y(\tau)\right) = \int_{\tau}^{\theta} S'_{\theta}\left(c^{*}, z\right) v'(\tau)\, \mathrm{d}z$$

has the same sign as $\theta - \tau$ since $S'_{\theta}\left(c^{*}, z\right) < 0$ for all z. It follows that (3) is satisfied for all τ and θ. This concludes the proof. □

Lemma 2 highlights a reversal of incentives following the introduction of random noise in redistribution. Indeed, by Assumption A1, the sub-utility $U(\theta)$ driving incentives turns decreasing with θ in the presence of noise ($v(\theta) > 0$ for some θ), though the overall utility $V(\theta)$ remains increasing with type ($V'(\theta) = u'_{\theta}\left(c^{*}, \theta\right) + \lambda S'_{\theta}\left(c^{*}, \theta\right) v(\theta) \simeq u'_{\theta}\left(c^{*}, \theta\right) > 0$ for λ close enough to 0 and $v(\theta)$ bounded from above).

The change in the direction of incentives captured by $U'(\theta)V'(\theta) \leq 0$ goes with a change in the sign of the informational rent $\lambda\left(S\left(c^{*}, \theta\right) - S\left(c^{*}, \tau\right)\right) v(\tau)$ given to type θ to avoid she mimics type τ. It now turns positive for $\tau > \theta$, much in contrast with the downward pattern that prevails when deterministic redistribution is used. In this sense, randomness in the redistribution policy aligns incentives with social preferences.

5 Welfare improving randomization

The argument put forward in favor of randomization in the literature is based on the relaxation of incentives when random noise is imposed on mimickees. It requires that these agents are less risk averse than the mimickers (Hellwig (2007)). The reversal of incentives in Lemma 2 is of a different nature since $S'_\theta(c^*,\theta) < 0$ is consistent with a coefficient of absolute risk aversion $A(c^*,\theta)$ that can be increasing (in Examples 1 and 2) as well as decreasing in θ (in Example 3). In the special case of logarithmic preferences considered in Example 3, where $A(c^*,\theta) = 1/(c^*+\theta)$, the alignment of social preferences and incentives can be achieved even though type θ^{inf} agents display the highest risk aversion.

Still, it is not clear at this stage whether random redistribution is socially useful since, by Lemma 2, incentive compatibility also requires a variance non-increasing with θ. That is, type θ^{inf} agents receive the greatest average transfers but also face the greatest noise. The following result gives a condition for the transfers to compensate for the loss from random taxation.

Proposition 1. Optimal random redistribution. *Consider a menu where the government uses random perturbations $\tilde{\varepsilon}(\theta)$ to the optimal deterministic after-tax income c^* with mean and variance $\lambda v(\theta)$ for some $\lambda \geq 0$ close to 0. Let $v(\theta)$ be non-negative, non-increasing and bounded from above. The random menu improves upon the deterministic optimum if and only if*

$$\int_{\Theta} \phi(c^*,\theta) v(\theta) \, \mathrm{d}F(\theta) > 0 \tag{9}$$

where

$$\phi(c,\theta) = S(c,\theta) - 1 - m(\theta) S'_\theta(c,\theta).$$

A proof is in Appendix A. A heuristic argument for deriving (9) follows from the perturbation methodology developed by Saez (2001), when applied to the total collected tax. Once random noise is introduced, the tax paid by a

θ agent is $y(\theta) - [c^* + \lambda v(\theta)]$. Using the definition of $U(\theta) = \lambda S(c^*, \theta) v(\theta) - y(\theta)$, the total collected tax can be written

$$\int\limits_{\Theta} [\lambda S(c^*, \theta) v(\theta) - U(\theta) - c^* - \lambda v(\theta)] \; \mathrm{d}F(\theta) \tag{10}$$

Consider a reform that increases the (rescaled) variance $v(\theta)$ of the after-tax income by a small amount $\mathrm{d}v$ for all types between θ and $\theta + \mathrm{d}\theta, \mathrm{d}\theta$ positive close to 0. These types, who are directly concerned by the reform, are in total number $f(\theta)\mathrm{d}\theta$. The argument distinguishes so-called behavioral and mechanical effects of the reform.

The behavioral effect is the change in collected tax abstracting from the adjustments needed to meet incentives, i.e., with $U(\theta)$ temporarily maintained fixed at its initial level. It captures the net social cost of the randomization that transits through the noise bearing on the socially favored agents. Every type θ directly concerned by the reform works more, which increases her before-tax income by $\lambda S(c^*, \theta) \mathrm{d}v$. The total tax resources thus increase by $\lambda S(c^*, \theta) f(\theta)\mathrm{d}v \ \mathrm{d}\theta$. By assumption, the government takes advantage of the noise to increase the average transfer to every such agents, which costs $\lambda \mathrm{d}v$ per agent. Overall the change in collected tax is $\lambda [S(c^*, \theta) - 1] f(\theta)\mathrm{d}v \ \mathrm{d}\theta$. Then, given $U(\theta)$, the reform yields a lower amount of collected taxes, which represents a net cost for the society (recall that $S(c^*, \theta) < 1$).

This cost has to be compared to the social gain from getting incentives aligned with redistribution tastes. It obtains when one accounts for the response of $U(\theta)$ to the reform that introduces random noise. This is the mechanical effect of the reform. By Lemma $2, U'(\theta)$ changes by $\mathrm{d}U'(\theta) = \lambda S'_\theta(c^*, \theta) \mathrm{d}v$ for every type directly concerned by the reform. It follows that the utility changes by $\mathrm{d}U = \mathrm{d}U'(\theta)\mathrm{d}\theta = \lambda S'_\theta(c^*, \theta) \mathrm{d}v \ \mathrm{d}\theta$ for every type above $\theta + \mathrm{d}\theta$, implying a change in total collected tax equal to $-(1 - F(\theta))\lambda S'_\theta(c^*, \theta) \mathrm{d}v \ \mathrm{d}\theta$. Since $S'_\theta \leq 0$, these indeed gives additional tax resources.

The final change in taxes following the introduction of the noise is

$$\lambda [S(c^*, \theta) - 1] f(\theta)\mathrm{d}v \ \mathrm{d}\theta - (1 - F(\theta))\lambda S'_\theta(c^*, \theta) \mathrm{d}v \ \mathrm{d}\theta$$

or equivalently,

$$\lambda [S(c^*, \theta) - 1 - m(\theta) S'_\theta(c^*, \theta)] f(\theta)\mathrm{d}v \ \mathrm{d}\theta.$$

The term into brackets is $\phi\left(c^{*},\theta\right)$ in Proposition 1. It balances the loss in taxes from agents concerned by the reform (their higher production does not compensate the cost from the additional transfers they receive) and the greater taxes allowed by the reduced informational rents given to high types above θ.

6 Shape of random redistribution

The inequality (9) shows that there is no social improvement from randomized taxes if $S'_{\theta}\left(c^{*},\theta\right)>0$, i.e., in the absence of the reversal of incentives (recall that $S\left(c^{*},\theta\right)<1$). It is satisfied if and only if $\phi(c^{*},\theta^{\mathrm{inf}})>0$ if the additional tax $\phi(c^{*},\theta)$ collected from type θ is decreasing with θ. In this case, there exists a threshold type θ^{*}, $\theta^{*}>\theta^{\mathrm{inf}}$, such that every type $\theta\leq\theta^{*}$ should face random taxes.

An extreme form of conflict between social redistribution and incentives occurs if $\phi(c^{*},\theta)$ instead is increasing in θ. Then a necessary, but not sufficient condition for (9), is $\phi(c^{*},\theta^{\mathrm{sup}})>0$. The government would like to design random taxes for high types specifically, but this would violate the monotonicity requirement for incentive compatibility that $v(\theta)$ must be non-increasing. Redistribution would then have to involve random taxes for all agents.

Exploiting the special form of (9), where the variance of the transfers enters linearly, allows us to go beyond these two polar configurations. To this purpose, we rely on the methodology introduced by Myerson (1981) and consider the new function

$$H(c^{*},q)=\int_{0}^{q}\phi(c^{*},F^{-1}(z))\,\mathrm{d}z$$

for every quantile $q\in[0,1]$ of the type distribution. Let G be the concave hull of H, and define $\bar{\phi}(c^{*},\theta)=G'_{q}(c^{*},F(\theta))$ as the so-called priority rule. We have:

Proposition 2. *Consider a menu where the government uses random perturbations $\tilde{\varepsilon}(\theta)$ to the optimal deterministic after-tax income c^{*} with mean and variance $\lambda v(\theta)$ for some $\lambda\geq 0$ close to 0. The random menu improves upon the deterministic optimum if and only if*

$$\bar{\phi}(c^{*},\theta^{\mathrm{inf}})>0.$$

There exists $\theta^* \geq \theta^{\inf}$ such that the highest amount of extra taxes from after-tax income randomization obtains by setting $v(\theta) > 0$ and non-increasing for all $\theta < \theta^*$, and $v(\theta) = 0$ for all $\theta \geq \theta^*$.

The result follows from the fact that $G'_q(c^*, F(\theta))$ is non-increasing in θ since $G(c^*, q)$ is concave in q. Its proof mirrors Myerson (1981), Section 6 pp. 68-69, or Condorelli (2012), and thus it is omitted. The two polar cases with $\phi(c^*, \theta)$ monotone in θ obtain for $H(q)$ either concave or convex for all θ. In the concave case, $\phi(c^*, \theta)$ is decreasing in θ, and $\bar{\phi}(c^*, \theta^{\inf}) = \phi(c^*, \theta^{\inf})$. In the convex case, $\phi(c^*, \theta)$ is increasing in θ,

$$\bar{\phi}(c^*, \theta) = \bar{\phi}(c^*, \theta^{\inf}) = H(c^*, 1) = \int_{\Theta} \phi(c^*, z)\, \mathrm{d}z$$

for all θ, and random taxes should be used if and only if a policy with uniform variance of taxes ($v(\theta) = v > 0$ for all θ) yields an extra amount of collected tax.

In view of Proposition 2, let us set $v(\theta) = v > 0$ if $\theta \leq \theta^*$, and $v(\theta) = 0$ otherwise. Then the inequality (9) is met if and only if there exists some threshold type $\theta^* \in \left[\theta^{\inf}, \theta^{\sup}\right]$ such that[2]

$$\frac{1 - S(c^*, \theta^{\inf})}{1 - S(c^*, \theta^*)} < 1 - F(\theta^*). \tag{11}$$

It cannot be that $\theta^* = \theta^{\inf}$ since this would imply that no agent faces income risk (indeed both sides of (11) then equal 1) while (9) is met. It can neither be that all agents should be exposed to random taxes: by Assumption 1, $S(c^*, \theta)$ is decreasing with θ and takes values in $[0, 1)$, the left-hand side of (11) is positive for all types while its right-hand side is 0 at $\theta^* = \theta^{\sup}$.

The writing (11) provides us with a better understanding of the economic conditions where random noise in redistribution can be useful. First, for the existence of an interior threshold where (11) is met, $S(c^*, \theta^{\inf})$ has to be much higher than $S(c^*, \theta)$ for almost all $\theta > \theta^{\inf}$, in which case the left-hand side of (11) gets closer to 0. Moreover, most agents should have high types close enough to $\theta^{\sup}$, so that $1 - F(\theta)$ in the right-hand side of (11) remains close to 1 for a large part of the population. Relying on the interpretation of θ as some given extra income or outside access capacities to housing, this

[2] A detailed derivation of (11) is in Appendix B.

corresponds to a situation where there are few poor, and thus a somewhat egalitarian distribution in this extra income or capabilities.

Overall, (11) shows that a combination of wide dispersion in the valuation of the good $S(c^*, \theta)$ and low variance in the extra income/capability θ favors redistribution by means of lotteries. Such a combination is reminiscent of circumstances identified by Weitzman (1977) and Spence (1977), where the price system is a better instrument than rationing to allocate resources. This may not come as a surprise. Indeed, when (11) is met, the socially favored types θ^{inf} value the good much more than the rest of the population, which gives the government strong incentives to leave them with a higher after-tax income or allocate a greater amount of goods. However, this cannot be achieved in a deterministic fashion, as deterministic fiscal tools yield bunching; all after-tax incomes are equal and some uniform rationing of consumption must be implemented. Intuitively, screening then is difficult to make since it has to be based on small differences in extra income/capabilities θ. Following Lemma 2, the government can instead base screening on valuation $S(c^*, \theta)$ in the presence of random noise. This allows, to some extent, for the replication of the market allocation, with higher, though stochastic, transfers to those actually rationed, who would have earned a higher after-tax income or consumed more through the market.

If applied to the case of housing, the deterministic optimum in the initial situation with bunching involves a form of uniform rationing, whereas the society would prefer to allocate more housing to agents with lower outside housing capabilities. By (11), randomness becomes socially beneficial in circumstances where the housing market would be superior to rationing, provided that the market would have led to allocations that satisfy the desired monotonicity. Related arguments may also apply to water distribution in developing countries threatened by water shortages and constrained to implement periodic rationing. If wealthy households have private pumping means that the public authorities cannot observe, the condition (11) delineates the circumstances in which one might consider shifting from uniform rationing to random cuts varying according to, e.g., the average wealth or property value of different neighborhoods. In the energy sector, where the wealthier may have private sources of supply (e.g., solar panels) that make public provision less attractive, some uncertainty in the public supply may be socially profitable as well.

7 Non-uniform partial bunching

So far we have considered the polar case of uniform bunching in the deterministic optimum where every type faces the same option (c^*, y^*). In practice, tax authorities often rely on tax schemes with more than two different tax brackets. Then, provided that bunching operates everywhere, the general form of the optimal deterministic tax schedule consists of a collection of income pairs (c_i^*, y_i^*) assigned to every agent with a type ranging from $\bar{\theta}_i$ to $\bar{\theta}_{i+1}$ $\left(\bar{\theta}_i < \bar{\theta}_{i+1}\right)$. Proposition 3 below gives a natural generalization of the uniform bunching alternative examined in Proposition 1 to such schedules with discontinuities.

Proposition 3. *Non-uniform deterministic bunching. Suppose that the optimal deterministic redistribution policy consists of n different tax brackets, with every type of agents in $\left[\bar{\theta}_i^*, \bar{\theta}_{i+1}^*\right)$ earning y_i^* before tax and c_i^* after-tax. There exists a random policy that improves upon the deterministic optimum if*

$$\sum_{i=1}^{n} \int_{\bar{\theta}_i^*}^{\bar{\theta}_{i+1}^*} \phi\left(c_i^*, \theta\right) \ v(\theta) \ \mathrm{d}F(\theta) > 0$$

for some non-increasing profile of after-tax income variance $(\mathrm{d}v(\theta))$ *close to* 0.

We outline the argument for the two-interval configuration $n = 2$ characterized by an interior threshold type $\bar{\theta}^*$ such that every type $\theta < \bar{\theta}^*$ earns c_1^* as after-tax income, while the remaining higher types $\theta \geq \bar{\theta}^*$ earns c_2^* $(c_1^* > c_2^*)$. This deterministic schedule is dominated if the introduction of small random noise $\mathrm{d}v(\theta) = \lambda v(\theta)$, $\lambda > 0$ small, on the after-tax income yields a higher amount of collected taxes while the socially favored type $\theta^{\inf}$ agents do not loose, $\mathrm{d}U(\theta^{\inf}) \geq 0$. The additional tax generated by the reform can then be redistributed to every agent through a uniform adjustment in before-tax income without violating incentive requirements.

The utility of every type $\theta < \bar{\theta}^*$ changes by

$$\mathrm{d}U(\theta) = \mathrm{d}U(\theta^{\inf}) + \int_{\theta^{\inf}}^{\theta} \lambda S_\theta' \left(c_1^*, z\right) \ \mathrm{d}v(z) \ \mathrm{d}z$$

so that the total change in utility from these agents can be written, after using the integration by parts formula,

$$\mathrm{d}U(\theta^{\inf}) - \left[1 - F\left(\bar{\theta}^*\right)\right] \mathrm{d}U\left(\bar{\theta}^*\right) + \int_{\theta^{\inf}}^{\bar{\theta}^*} \lambda S'_\theta\left(c_1^*, \theta\right) m(\theta)\ \mathrm{d}v(\theta)\ \mathrm{d}z\ \mathrm{d}F(\theta).$$

Similarly, the utility of every type $\theta \geq \bar{\theta}^*$ changes by

$$\mathrm{d}U(\theta) = \mathrm{d}U(\theta^{\inf}) + \int_{\theta^{\inf}}^{\bar{\theta}^*} \lambda S'_\theta\left(c_1^*, z\right)\ \mathrm{d}v(z)\ \mathrm{d}z + \int_{\bar{\theta}^*}^{\theta} \lambda S'_\theta\left(c_2^*, z\right)\ \mathrm{d}v(z)\ \mathrm{d}z,$$

which now yields a total utility change for these agents equal to

$$\left[1 - F\left(\bar{\theta}^*\right)\right] \mathrm{d}U\left(\bar{\theta}^*\right) + \int_{\bar{\theta}^*}^{\theta^{\sup}} \lambda S'_\theta\left(c_2^*, \theta\right) m(\theta)\ \mathrm{d}v(\theta)\ \mathrm{d}F(\theta).$$

Then, from (10), the additional amount of collected tax implied by the introduction of random noise writes

$$\int_{\theta^{\inf}}^{\bar{\theta}^*} \lambda\phi\left(c_1^*, \theta\right)\ \mathrm{d}v(\theta)\ \mathrm{d}F(\theta) + \int_{\bar{\theta}^*}^{\theta^{\sup}} \lambda\phi\left(c_2^*, \theta\right)\ \mathrm{d}v(\theta)\ \mathrm{d}F(\theta) - \mathrm{d}U(\theta^{\inf}).$$

Given a profile $(\mathrm{d}v(\theta))$, which Lemma 2 shows must be non-increasing to meet incentive requirements, the highest amount of collected tax that does not hurt type $\theta^{\inf}$ agents obtains by setting $\mathrm{d}U\left(\theta^{\inf}\right) = 0$. This yields the condition given in Proposition 3 for $n = 3$, with $\bar{\theta}_1^* = \theta^{\inf}, \bar{\theta}_2^* = \bar{\theta}^*$ and $\bar{\theta}_3^* = \theta^{\sup}$.

Remark 1. Partial bunching. Proposition 3 also applies for $\bar{\theta}_{n+1}^* < \theta^{\sup}$, i.e., in the absence of bunching at the top of the distribution. Then, one can set $v(\theta) = 0$ for all types that are not concerned by bunching, $\theta \geq \bar{\theta}_{n+1}^*$. The change in collected tax is

$$\sum_{i=1}^{n} \int_{\bar{\theta}_i^*}^{\bar{\theta}_{i+1}^*} \lambda\phi\left(c_i^*, \theta\right) \mathrm{d}v(\theta) \mathrm{d}F(\theta) - \mathrm{d}U(\theta^{\inf}) + \left[1 - F(\bar{\theta}_{n+1}^*)\right] \mathrm{d}U(\bar{\theta}_{n+1}^*).$$

As above, $\mathrm{d}U(\theta^{\inf}) = 0$ maximizes the additional tax revenue. The perturbation argument guarantees incentive compatibility among types below $\bar{\theta}^*_{n+1}$. To avoid failures of incentives involving types above $\bar{\theta}^*_{n+1}$ one can give $\mathrm{d}U(\bar{\theta}^*_{n+1})$ to every such type above $\bar{\theta}^*_{n+1}$. This costs $[1 - F(\bar{\theta}^*_{n+1})]\mathrm{d}U(\bar{\theta}^*_{n+1})$ in terms of tax resources, hence the result in Proposition 3 for this special case.

8 A parametric example

We now exhibit a specific parametrization of the economy where bunching occurs for deterministic fiscal tools and tax randomization is socially useful.

Preferences are represented by the logarithmic utility function $u(c, \theta) = \ln(c + \theta)$ used in Example 3. The parameter θ is distributed according to a generalized Weibull distribution.

Using (6) the optimal deterministic relaxed redistribution policy maximizes

$$V\left(\theta^{\inf}\right) = \int_{\Theta} \left[\ln(c(\theta) + \theta) - c(\theta) - \frac{m(\theta)}{c(\theta) + \theta}\right] \mathrm{d}F(\theta). \tag{12}$$

The after-tax income $c^*(\theta)$ that maximizes pointwise $V(\theta^{\inf})$ is a nonnegative root of the first-order condition $(c^*(\theta) + \theta)^2 - (c^*(\theta) + \theta) - m(\theta) = 0$. There is only one possible such root,

$$c^*(\theta) = \frac{1 + (1 + 4m(\theta))^{1/2}}{2} - \theta. \tag{13}$$

The optimal deterministic after-tax income is $c^*(\theta)$ if this quantity is nonincreasing. Otherwise, i.e., if $c^*(\theta)$ is increasing, there is bunching in the deterministic case. Thus bunching occurs occurs if and only if

$$m'(\theta) > (1 + 4m(\theta))^{1/2}. \tag{14}$$

The above inequality cannot be met for standard probability distributions, as they have a decreasing Mills ratio, $m'(\theta) \leq 0$. It may however be satisfied for well-chosen log-logistic, Weibull, and variants of Weibull distributions such as generalized or power generalized Weibull commonly used in econometric models for duration data. Where it holds true, $c^*(\theta)$ violates the second-order monotonicity requirements for incentive compatibility. Our example uses a generalized Weibull distribution (see Dimitrakopoulou, Adamidis, and

Loukas (2007) for properties of this distribution). Its cumulative distribution function is

$$F(\theta) = 1 - \exp\left[1 - (1 + \lambda\theta^b)^a\right]$$

for $\theta \geq 0$, with a, b and λ positive parameters. The Mills ratio $m(\theta)$ is increasing for $a < 1$ and $b \leq 1$.

We set $a = 0.5, b = 0.05$ and $s = 0.5$. The condition (14) for bunching in the optimal deterministic policy is satisfied if and only if $\theta \leq 19.9$, which corresponds to 22.7 percent of the population with the lowest types.

The optimal deterministic income tax schedule consists of a single income pair $(\bar{c}^*, \bar{y}^*)$ offered to every type $\theta \leq \bar{\theta}^*$ while all the other types are assigned the optimal relaxed income pair $(c^*(\theta), y^*(\theta))$. The social objective $V(\theta^{\text{inf}})$ thus is

$$\int_0^{\bar{\theta}^*} \left[\ln(\bar{c}^* + \theta) - \bar{c}^* - \frac{m(\theta)}{\bar{c}^* + \theta}\right] \mathrm{d}F(\theta) \tag{15}$$

$$+ \int_{\bar{\theta}^*}^{+\infty} \left[\ln(c^*(\theta) + \theta) - c^*(\theta) - \frac{m(\theta)}{c^*(\theta) + \theta}\right] \mathrm{d}F(\theta).$$

Using the expression of $c^*(\theta)$ given in (13), the optimal threshold is such that

$$\bar{c}^* = c^*\left(\bar{\theta}^*\right) = \frac{1 + \left(1 + 4m\left(\bar{\theta}^*\right)\right)^{1/2}}{2} - \bar{\theta}^* \tag{16}$$

To characterize the amount of after-tax income $\bar{c}^*$, we apply the integration by parts formula and rewrite the contribution of types below $\bar{\theta}^*$ to the social objective $V(\theta^{\text{inf}})$ in (15) as

$$-\left[1 - F\left(\bar{\theta}^*\right)\right] \ln\left(\bar{c}^* + \bar{\theta}^*\right) + \ln \bar{c}^* - \bar{c}^* F\left(\bar{\theta}^*\right).$$

Solving for the first-order condition for $\bar{c}^*$ to maximize this contribution, which is a quadratic equation in $\bar{c}^*$, the only positive root is

$$\bar{c}^* = \frac{1 - \bar{\theta}^*}{2} + \frac{1}{2}\sqrt{\left(1 - \bar{\theta}^*\right)^2 + \frac{4\bar{\theta}^*}{F(\bar{\theta}^*)}}.$$

Replacing this expression of $\bar{c}^*$ into (16) defines the optimal threshold $\bar{\theta}^*$. Numerical computations (see the `R` code in Appendix B) yield $\bar{c}^* = 4.02$ and $\bar{\theta}^* = 74.87$, with $F(\bar{\theta}^*) = 23.88$ percent.

Following Remark 1, we now expose to random noise a subset of low types among those facing the after-tax income $\bar{c}^*$ in the deterministic income tax schedule. Namely we set $\mathrm{d}v(\theta) = \mathrm{d}v > 0$ for all $\theta \leq \bar{\theta}$, with $\bar{\theta}$ some threshold type below $\bar{\theta}^*$, while $\mathrm{d}v(\theta) = 0$ for all $\theta > \bar{\theta}$. With this variance step-profile, the condition for socially useful randomness in redistribution given in Proposition 3 can be re-expressed as

$$-F(\bar{\theta}) - \frac{1-F(\bar{\theta})}{\bar{c}^* + \bar{\theta}} \left(1 - \frac{1}{2\left(\bar{c}^* + \bar{\theta}\right)}\right) + \frac{1}{\bar{c}^*}\left(1 - \frac{1}{2\bar{c}^*}\right) > 0. \tag{17}$$

The shape of the left-hand side is depicted in Figure 1. It is 0 when evaluated at $\bar{\theta} = 0$ and decreasing for $\bar{\theta}$ close enough to 0, so that one should not set random taxes on a too narrow subset of types close to $\theta^{\inf}$. For higher values of $\bar{\theta}$, it is single-peaked, reaching its global maximum of 0.0265 for $\bar{\theta} = 73.83$, hence very close to $\bar{\theta}^* = 74.87$. It takes positive values for all $\bar{\theta} \in \left[9.76, \bar{\theta}^*\right]$, with $F(9.76) = 7.33$ percent.

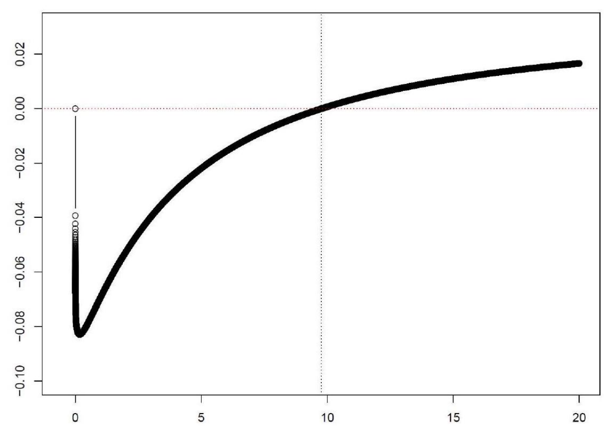

Figure 1: Random redistribution with a generalized Weibull distribution

The figure depicts function of $\bar{\theta}$ that appears in the left-hand side of (17). It is drawn for a generalized Weibull distribution with parameters $a = 0.5, b = 0.05$ and $s = 0.5$. The threshold $\bar{\theta}$ is on the horizontal axis. See the R *code in Appendix B for recovering the figure. It is optimal to expose all agents with type θ below $\bar{\theta}$ if the function takes a positive value when evaluated at $\bar{\theta}$. The vertical dotted line at $\bar{\theta} = 9.76$ gives the least value of the threshold such that the function reaches positive values. For readability purposes, the figure does not represent the function for $\bar{\theta}$ above 20. The function is actually single-peaked at takes positive values for $\bar{\theta}$ below $\bar{\theta}^* = 74.87$, and negative values for higher $\bar{\theta}$, so that it is not optimal to expose all agents to random taxation.*

We conclude that, in this example, introducing small random perturbations on taxes designed for the bottom of the type distribution improves upon the deterministic optimum. Namely, randomness should concern between at least the bottom 7.33 and at most 23.88 percent of types. Relying on the interpretation of (9) as a change in collected tax following the introduction of income risk, the highest amount of additional taxes would be generated by subjecting almost all agents affected by bunching to random perturbations, but not all of them. The highest social welfare gain that can be achieved equals $0.0265 \times \lambda v$ income units. Since the less well-off get at most $\ln(\bar{c}^* + \theta^{\inf}) = \ln(4.02)$ units of before-tax income, this gain represents a $0.0265/\ln(4.02) \times \lambda v$ share of the initial level of welfare. For $\lambda v = 1$, i.e., a one-unit increase in the average transfer (25 percent of the initial after-tax income), we find a modest welfare gain, with lower bound of 2 percent.

9 Conclusion

Our paper reexamines optimal income taxation in a Mirrlees setup with a continuum of types of taxpayers. Our focus is on the choice between deterministic versus random taxation. We have shown that the random alternative is preferred only if the best deterministic policy implies a uniform treatment of different types of taxpayers. Randomness then allows the government to exploit taxpayers' risk aversion and implement discriminatory tax treatment.

The existing literature following Hellwig (2007) suggests that rationing, viewed as implying randomness in the allocation of goods designed for the poor, can be justified as far as these agents display lower risk aversions. However we expect a greater, not lower aversion to consumption risk among the poor, those who have the highest marginal utility gain from consumption. Our paper shows that stochastic redistribution can be socially useful even though random noise bears on the most risk averse agents. In this respect, it can be used to justify policies relying on rationing the less well-off part of the population to improve its welfare. This may be especially relevant in the case of housing or in the provision of public goods when recipients differ in the availability of substitutes that are difficult to observe.[3]

[3] Our framework can accommodate applications to minimum wage and other forms of rationing in the labor market discussed in Introduction by considering utilities linear in after-tax, rather than before-tax income as in (1). Random noise then is on labor and before-tax income.

In a more narrowly fiscal perspective, the social acceptability of explicit forms of randomization of taxes in the tax code is plausibly disputable. Non-explicit randomization from administrative errors because of imprecise assessment of before-tax income as in Stern (1982) or Slemrod (2019), make a deterministic tax code consistent with small random income perturbations. Actually, it is likely that such non-explicit forms of randomization bear on the less well-off part of the population, as the most vulnerable usually fall into several social benefit regimes. In the redistributive case, where these agents have high risk aversion and high social importance, our results suggest that it may be wasteful to correct these errors.

Two features in our analysis would be worth addressing in further work. First, we considered the case of a Rawlsian planner, which magnifies tensions from redistribution. A continuity argument suggests that the results should remain unaffected for weighted utilitarian redistributive preferences that place greater importance on agents who value consumption more. On the other hand, the occurrence of bunching in the deterministic optimum may be less plausible for weak redistribution motives, e.g., the unweighted (Benthamite) utilitarian social welfare objective, implying low redistribution made deterministically.

A second feature relates to the interplay between the extent of bunching and optimal randomization. In our parametric example in Section 8, random taxation applies to a subset of the agents affected by failures of monotonicity requirements while redistribution should remain deterministic at the top of the distribution. It may thus be that deterministic redistribution is more suitable when a smaller portion of the population is affected by bunching.

References

Abajian, A., C. Cole, B.K. Jack, K. Meng, and M. Visser (2024): "Dodging Day Zero: Drought, adaptation, and inequality in Cape Town," Working Paper.

Brito, D., J. Hamilton, S. Slutsky, and J. Stiglitz (1995): "Randomization in optimal income tax schedules," Journal of Public Economics, 56, 189-223.

Chetty, R. (2006): "A new method of estimating risk aversion," American Economic Review, 96 (5): 1821-1834.

Cohen, A. and L. Einav (2007): "Estimating risk preferences from deductible choice," American Economic Review 97, 745-788.

Dimitrakopoulou, T., K. Adamidis, and S. Loukas (2007): "A Lifetime Distribution With an Upside-Down Bathtub-Shaped Hazard Function," IEEE Transactions on Reliability, 56(2), 308-311.

Dworczak, P., S. D. Kominers, and M. Akbarpour (2021): "Redistribution through markets," Econometrica, 89(4), 1665-1698.

Ederer, F., R. Holden, and M. Meyer (2018): "Gaming and strategic opacity in incentive provision," The RAND Journal of Economics, 49(4), 819-854.

Gauthier, S., and G. Laroque (2014): "On the value of randomization," Journal of Economic Theory, 151, 493-507.

Gauthier, S., and G. Laroque (2017): "Redistribution by means of lotteries," Journal of Economic Theory, 169, 707-716.

Hellwig, M. (2007): "The undesirability of randomized income taxation under decreasing risk aversion," Journal of Public Economics, 91, 791-816.

Laffont, J.-J., and D. Martimort (2002): The theory of incentives: The principalagent model, Princeton University Press.

Lang, M. (2017): "Legal Uncertainty as a welfare enhancing screen," European Economic Review, 91, 274-289.

Lang, M. (2023): "Stochastic Contracts and Subjective Evaluations," Rand Journal of Economics, 54, 104-134.

Lollivier, S., and J.-C. Rochet (1983): "Bunching and Second-Order Conditions: A Note on Optimal Tax Theory," Journal of Economic Theory, 31, 392-400.

Mirrlees, J. (1971): "An exploration in the theory of optimum income taxation," Review of Economic Studies, 38, 175-208.

Myerson, R. (1981): "Optimal Auction Design," Mathematics of Operations Research, 6, 58-73.

Nichols, A. L., and R. J. Zeckhauser (1982): "Targeting Transfers through Restrictions on Recipients," American Economic Review, 72, 372-377.

Pavlov, G. (2011): "Optimal mechanism for selling two goods," The B.E. Journal of Theoretical Economics, 11, 1-35.

Pycia, M., and U. Unver (2015): "Decomposing random mechanisms," Journal of Mathematical Economics, 61, 21-33.

Saez, E. (2001): "Using Elasticities to Derive Optimal Income Tax Rates," Review of Economic Studies, 68, 205-229.

Slemrod, J. (2019): "Tax Compliance and Enforcement," Journal of Economic Literature 57, 904-954.

Spence, M., (1977): "Nonlinear prices and welfare," Journal of Public Economics, 8, 1-18.

Stern, N.H. (1982): "Optimum taxation with errors in administration", Journal of Public Economics, 17, 181-211.

Stiglitz, J. (1981): "The Allocation Role of the Stock Market," The Journal of Finance, 36, 235-251.

Stiglitz, J. (1982): "Utilitarianism and horizontal equity: the case for random taxation," Journal of Public Economics, 18, 1-33.

Strausz, R. (2006): "Deterministic versus stochastic mechanisms in principal-agent models," Journal of Economic Theory, 128, 306-314.

Tobin, J. (1971): "On limiting the domain of inequality," Journal of Law and Economics, 13, 263-277.

Weiss, L. (1976): "The desirability of cheating incentives and randomness in the optimal income tax," Journal of Political Economy, 84, 1343-1352.

Weitzman, M. (1977): "Is the price system or rationing more effective in getting a commodity to those who need it most?," Bell Journal of Economics, 8, 517-524.

Wijkander, H. (1988): "Equity and efficiency in public sector pricing: A case of stochastic rationing," Econometrica, 56, 1455-1465.

Appendices

A A Proof of Proposition 1

We first express the before-tax income of type θ as a function of her indirect utility $V(\theta)$,

$$y(\theta) = u(c^*, \theta) + \lambda S(c^*, \theta)\, v(\theta) - V(\theta).$$

The expression of the indirect utility $V(\theta)$ obtains from using the first-order necessary condition in Lemma 2 for incentive compatibility, which yields

$$U(\theta) = U\left(\theta^{\inf}\right) + \int_{\theta^{\inf}}^{\theta} \lambda S'_\theta(c^*, z)\, v(z)\mathrm{d}z,$$

so that

$$V(\theta) = u(c^*, \theta) + U\left(\theta^{\inf}\right) + \int_{\theta^{\inf}}^{\theta} \lambda S'_\theta(c^*, z)\, v(z)\mathrm{d}z.$$

The feasibility constraint (2) reads

$$\int_{\Theta} \left[c^* + \lambda v(\theta) - y(\theta)\right] \mathrm{d}F(\theta) = 0.$$

Replacing the before-tax income $y(\theta)$ with its expression in terms of $V(\theta)$, with $V(\theta)$ given above, we find

$$\int_{\Theta} \left[c^* + \lambda v(\theta) - \lambda S(c^*, \theta)\, v(\theta)\right] \, \mathrm{d}F(\theta)$$

$$+U\left(\theta^{\inf}\right) + \int_{\Theta} \int_{\theta^{\inf}}^{\theta} \lambda S'_\theta(c^*, z)\, v(z)\, \mathrm{d}z\, \mathrm{d}F(\theta) = 0.$$

Using the integration by parts formula,

$$\int_{\Theta} \int_{\theta^{\inf}}^{\theta} S'_\theta(c^*, z)\, v(z)\, \mathrm{d}z\, \mathrm{d}F(\theta) = \int_{\Theta} m(\theta) S'_\theta(c^*, \theta)\, v(\theta)\, \mathrm{d}F(\theta),$$

the feasibility constraint allows us to get the sub-utility $U(\theta)$ driving incentives in the presence of small random tax perturbations for type $\theta^{\inf}$,

$$U\left(\theta^{\inf}\right) = -\int_{\Theta} \left[c^* + \lambda v(\theta) - \lambda S\left(c^*, \theta\right) v(\theta) + m(\theta)\lambda S'_\theta\left(c^*, \theta\right) v(\theta)\right] \mathrm{d}F(\theta).$$

Social welfare is $V\left(\theta^{\inf}\right) = u\left(c^*, \theta^{\inf}\right) + U\left(\theta^{\inf}\right)$, which is actually

$$u\left(c^*, \theta^{\inf}\right) - \int_{\Theta} \left[c^* + \lambda\phi\left(c^*, \theta\right) v(\theta)\right] \mathrm{d}F(\theta),$$

with $\phi\left(c^*, \theta\right)$ defined in Proposition 1.

The expression of social welfare in the absence of noise obtains by letting $v(\theta) = 0$ for all θ. It reduces to $u\left(c^*, \theta^{\inf}\right) - c^*$. This yields condition (9) in Proposition 1 for socially useful random redistribution (recall that $\lambda \geq 0$ for the variance of the income to be non-negative). This concludes the proof.

B Detailed derivation of (11)

Applying the integration by parts formula, we have

$$\int_{\theta^{\inf}}^{\theta^*} m(\theta) S'_\theta(c^*, \theta)\, \mathrm{d}F(\theta) = \left[1 - F(\theta^*)\right] S(c^*, \theta^*) - S(c^*, \theta^{\inf}) + \int_{\theta^{\inf}}^{\theta^*} S(c^*, \theta)\, \mathrm{d}F(\theta).$$

Therefore,

$$\int_{\theta^{\inf}}^{\theta^*} \phi(c^*, \theta)\, \mathrm{d}F(\theta) = -F(\theta^*) - \left[1 - F(\theta^*)\right] S(c^*, \theta^*) + S(c^*, \theta^{\inf}).$$

This is positive if and only if

$$S(c^*, \theta^{\inf}) - S(c^*, \theta^*) > F(\theta^*)\left[1 - S(c^*, \theta^*)\right].$$

Since $S(c^*, \theta^*) < 1$, this rewrites as

$$\frac{S(c^*, \theta^{\inf}) - S(c^*, \theta^*)}{1 - S(c^*, \theta^*)} > F(\theta^*),$$

which is equivalent to (11).

C R code for Section 8

```
a <- 0.5; b <- 0.05; s <- 0.5
  FF <- function(x) 1 - exp (1-(1+s*x^b)^a)
  ff <- function(x) {
    ff <- (a*(1+s*x^b )^(a-1)*s*b*x^(b-1))
    ff <- ff*exp(1-(1+s*x^b)^a)
    ff
  }
  mm <- function(x) (1-FF(x)) / ff(x)
  mmprime <- function(x) {
    mmp <- - (a*b*(b-1)*s*x^(b-2)*(1+s*x^b)^(a-1))
    temp <- a*b*s*x^(b-1)*(a-1)*s*b*x^(b-1)
    mmp <- mmp - temp*( 1+s*x^b)^(a-2)
    mmp <- mmp / (a*b*s*x^(b-1)*( 1+s*x^b)^(a-1))^2
    mmp
  }

bunch <- function(x) mmprime(x) - (1+4*mm(x))^(1/2)
  xx <- seq (1e-10,1e3,1e-2)
  plot (xx , bunch(xx))
        # bunching occurs for xx such that bunch (xx) is positive
  max (xx[bunch(xx)>=0]); FF(max(xx[bunch(xx)>=0]))

thetabar <- function(x) x+((1-x)^2+4*x/FF(x))^(1/2)-(1+4*mm(x))^(1/2)
  plot (xx, thetabar(xx))
        # threshold below which bunching occurs has thetabar = 0
  min(xx[thetabar(xx)>=0]); FF(min(xx[thetabar(xx)>=0]))
  theta <- min(xx[thetabar(xx)>=0])
  theta

cbar <- (1-theta)/2+((1-theta)^2+4*theta/FF(theta))^(1/2)/2
  cbar

sfnum <- function(x) {
  sf <- (1-FF(x))*(1-1/(2*(cbar+x)))/(cbar+x)
  sf <- -FF(x)-sf+(1-1/(2*cbar))/cbar
  sf
  }
```

```
  sumphi <- function(x) sfnum(x)-sfnum(1e-10) # sumphi is 0 at 1e-10
  xx <- seq (1e-10, theta, 1e-1)
  max(sumphi(xx)); xx[sumphi(xx)==max(sumphi(xx))]
  sumphi(theta)
  min(xx[sumphi(xx)>0]); FF(min(xx[(xx)>0]))
# Figure exported in the main text
xx <- seq(1e-10, 2, 1e-5)
  plot(xx, sumphi(xx), type="b", xlim =c(-0.5,20), ylim =c(-0.1,0.03))
  xx <- seq(2,20,1e-3)
  points(xx, sumphi(xx), type ="b")
  abline(h=0, col="red", lty="dotted")
  abline(v= min(xx[sumphi(xx)>0]), lty ="dotted")
```