Attrition in Randomized Control Trials: Using tracking information to correct bias

Teresa Molina-Millán and Karen Macours*

January 21, 2021

*Molina-Millán: University of Alicante, (teresa.molina@ua.es); Macours: Paris School of Economics and INRAE,(karen.macours@psemail.eu). Acknowledgments: This research would not have been possible without the support of Ferdinando Regalia of the Inter-American Development Bank (IDB), and co-PIs Tania Barham and John Maluccio in the wider research project. We gratefully acknowledge generous financial support from IDB, the Initiative for International Impact Evaluation (3ie: OW2.216), the National Science Foundation (SES 11239945 and 1123993) and the French National Research Agency (ANR), under Grant ANR-17-EURE-0001. We are indebted to Veronica Aguilera, Enoe Moncada, and the survey team from CIERUNIC for excellent data collection and for their dogged persistence in tracking. We are grateful for many related discussions and ideas from Tania Barham, Luc Behaghel, John Maluccio, Joachim De Weerdt as well as comments received during presentations at Paris School of Economics, the Inter-American Development Bank, and LACEA, and from two anonymous referees. All remaining errors and omissions are our own.

Abstract

This paper starts from a review of RCT studies in development economics, and documents many studies largely ignore attrition once attrition rates are found balanced between treatment arms. The paper analyzes the implications of attrition for the internal and external validity of the results of a randomized experiment with balanced attrition rates, and proposes a new method to correct for attrition bias. We rely on a 10-years longitudinal data set with a final attrition rate of 10 percent, obtained after intensive tracking of migrants, and document the sensitivity of ITT estimates for schooling gains and labor market outcomes for a social program in Nicaragua. We find that not including those found during the intensive tracking leads to an overestimate of the ITT effects for the target population by more than 44 percent, and that selection into attrition is driven by observable baseline characteristics. We propose to correct for attrition using inverse probability weighting with estimates of weights that exploit the similarities between missing individuals and those found during an intensive tracking phase. We compare these estimates with alternative strategies using regression adjustment, standard weights, bounds or proxy information.

1 Introduction

Longitudinal surveys and panel datasets are indispensable tools for the study of economic and demographic dynamics in developing countries. Randomized Control Trials (RCTs) in particular rely on panel datasets to estimate impacts of randomized interventions on the target population. Keeping a panel dataset representative of the study population requires addressing attrition. Potential biases resulting from missing individuals across waves in panel surveys have long been a focus of study for longitudinal data in developed countries,[1] but the literature on developing countries is more limited. The many data collection efforts in the context of RCT studies in developing countries, however, point to the relevance of better understanding attrition in such contexts. This is even more so as more analyses start focusing on the longer-term effects and dynamics after RCTs.

Attrition in developing countries is often driven by migration, and the decision to migrate or the correlates of migration might well be affected by the randomized exposure to particular interventions. The challenges posed by attrition therefore could be different than those in developed countries where it is often related to refusals. While a number of longitudinal studies in developing countries document that those who are missing differ in observable characteristics from those who are found, (Alderman et al. (2001);Thomas, Frankenberg and Smith (2001); Falaris (2003); Baird, Hamory and Miguel (2008);Thomas et al. (2012)) the implications for Intent-to-treat (ITT) estimations in impact evaluations are not always fully taken into account. Apart from the reduction in the number of observations, and the related loss of statistical power, attrition can reduce internal validity in case it leads to unbalanced samples. It can also have important implications for external validity in the presence of heterogeneous treatment effects, if treatment effects are different for attritors than for the rest of the study population. In addition, take-up of an intervention is likely to be lower for people that migrate prior or during an intervention. Omitting migrants can then lead to an overestimate of the ITT effects. Data collection to estimate impacts of RCTs, however, often does not include protocols to track migrants

[1] See, for instance, issue number 33 of the Journal of Human Resources (1998) on "Attrition in Longitudinal surveys".

outside their village or community of origin, sometimes resulting in high attrition rates. Given the mobility of many target populations migrant tracking can be costly, and data collection cost-concerns often need to be weighted against the potential consequences of selection bias due to attrition.

This paper exploits different phases of the tracking protocol of a longitudinal impact evaluation survey to illustrate the potential challenges resulting from non-random attrition in RCTs. We first show how commonly used tracking protocols would have led to an overestimation of the treatment effect for the population under study, and then show how information from different stages of the tracking process can be used to account for the remaining attrition.

To motivate the analysis, we review how RCT studies in development economics handle attrition.[2] Survey attrition rates vary widely even for similar target populations. Average annual attrition rates in studies targeting respondents below 18 years old, for instance, vary from 0 to 60 percent. Notably, the consideration of the potential attrition bias is often limited, in contrast with the care given ex-ante to assure random program placement. Around 23 percent of studies do not go beyond testing whether attrition rates between treatment arms are different, and 18 percent of studies do not even show such test. For studies that address attrition in more detail, a variety of approaches can be found, including studying the correlates of attrition (44%), showing balance after attrition (28%) or analyzing treatment heterogeneous effects on attrition (45%). Overall, only 28% of studies explicitly corrects for attrition in the estimations, with non-parametric bounds and Inverse Probability Weights (IPW) the most common methodologies applied. Among studies in which the authors identified non-random attrition, around 31 percent do not apply a sample-selection correction method to correct for attrition. On the other hand, around one third percent of studies showing balanced attrition rates still apply a sample-selection correction method.[3]

To quantify the implications of different approaches to attrition for a specific case,

[2]See Appendix G.2 for details on the selection of papers and the different findings.

[3]While addressing attrition through methods beyond balance tests has become more common in recent literature, qualitatively the summary of practices in the literature is not very different when only considering more recently published work. See Table G4 and G6.

this paper analyzes the incidence and implications of attrition on a 10 years longitudinal data set, collected for a randomized evaluation of a Conditional Cash Transfer (CCT) implemented from 2000 to 2005 in Nicaragua. We use data from a pre-program census collected in 2000 and from a follow-up survey conducted between November 2009 and November 2011. Barham, Macours and Maluccio (2017) use this data to estimate the 10-year after impacts of the CCT program. Considerable effort was made during the tracking process of the follow-up sample to reduce attrition and to interview permanent and temporal migrants. The tracking process lasted almost 2 years and individuals were followed everywhere in Nicaragua and Costa Rica, the destination country for the vast majority of international migrants from the study population. We distinguish between a *Regular Tracking Phase* (RTP) covering all localities included in the original survey sample, and an *Intensive Tracking Phase* (ITP) in which individuals that could not be located during the RTP were tracked intensively. The division between regular and intensive tracking corresponds to normal and high-effort tracking process, where the regular tracking process is similar to the common protocol in many surveys. Attrition was almost 30 percent after the RTP, similar to attrition rates also found in young mobile population in other studies (such as the 10-year evaluation of the related CCT program in Mexico (Behrman, Parker and Todd, 2009)). Attrition falls to 10 percent after the ITP and the data collected during the ITP allows quantifying the attrition bias obtained after regular tracking only.

We first show that response rates are balanced between treatment arms at different stages during the tracking process. We then analyze the implications of attrition by estimating ITT effects for different subsamples corresponding to the different phases of the tracking process. We estimate the ITT coefficient of the CCT on two long-term outcomes of the program, long-term gain in grades of education attained and off-farm employment of boys aged between 9-12 at the start of the program.[4] The ITT estimates suggest that

[4]The grades of education attained is the direct long-term outcome of the CCT program, which had as one of its main objectives increasing school attainment. The off-farm employment outcome can be seen as a targeted final outcome of the intervention, consistent with the CCTs objective to increase human capital in order to improve ex-beneficiaries' long-term economic outcomes in the labor market. As the two outcome variables can conceptually be seen as causally related, one could hypothesize attrition bias to go in the same direction. However, if one hypothesizes that education gains mostly occurs in villages

the CCT increased schooling by 0.61 years (p-value 0.01) for boys found using the regular tracking procedures (RTP). The estimate using the whole sample of boys surveyed in 2010 is almost one third lower (0.43 years) than the RTP estimate (significantly different at the 5 percent), suggesting that without conducting an intensive tracking protocol we would have overestimate the ITT estimate on the change in years of schooling. A similar pattern is observed when we estimate the ITT coefficient on off-farm employment, with the ITT coefficient after regular tracking being 9 percentage points, compared to 6 percentage points with the full sample. The ITT estimates are also sensitive- and indeed further decrease- when controlling for additional baseline variables and more so after regular tracking than on the full sample. These findings can be explained by analyzing the correlates of attrition, as we find that attrition is correlated with many baseline observables, capturing socio-economic status, demographic composition of the household, family networks and the potential temporary nature of the baseline residence. Moreover, these characteristics relate differently to attrition in the two experimental groups, indicating that this may well lead to bias in the ITT estimates.

A comparison of baseline characteristics by the respondent's status at the end of the follow-up survey (found during RTP, found during ITP, never found) shows that those who were never found are relatively more similar in baseline characteristics to those in the ITP sample than to those in the RTP sample. Thus, we propose a new method to correct for attrition bias exploiting the similarities, in observable characteristics, between attritors and the intensive tracking sample. We build on Fitzgerald, Gottschalk and Moffitt (1998) and Wooldridge (2002*b*) and estimate the probability to be found to construct Inverse Probability Weights. But instead of using baseline information for the complete sample of respondents in the follow-up, we estimate weights using only information on the sample of respondents tracked during the ITP. The underlying assumption is that those found and not-found in the ITP sample are more similar in both observed and unobserved characteristics, than those in the *Complete Tracking Phase* (CTP) sample. We further

of origin, that increases in education do not necessarily lead to more migration, but that migrants are more likely to be those with off-farm employment, one could also expect attrition bias to work differently for both outcomes.

show that the observed characteristics have more explanatory power in the ITP sample than in the CTP sample. Estimates with the new IPW lead to smaller point estimates than the ITT, and are more robust to different specifications of the control variables. In contrast, applying standard IPW to the full sample or the regular tracking sample leads to less robust estimates.

We also apply other common approaches to account for attrition and use the information of the intensive tracking to assess the plausibility of the assumptions underlying the different estimates. Assuming worst-case scenario (Horowitz and Manski, 2000) to calculate bounds leads to large and uninformative bounds for both outcomes, and the same holds when using more stringent assumption about the non-respondents following Kling, Liebman and Katz (2007). In contrast Lee bounds (Lee, 2002, 2009) after both regular and intensive tracking lead to intervals that do not include the new IPW benchmark estimate. Analysis of the correlates of attrition further suggests that the monotonicity assumption for the Lee Bounds may not hold in this context.

This paper relates to the econometric literature on sample selection.[5] It proposes an alternative approach, building on Fitzgerald, Gottschalk and Moffitt (1998) methodology of constructing a model specific IPW, that exploits similarities between difficult-to-find respondents and attritors to correct for attrition bias. It shares with Behaghel et al. (2015)'s selectivity correction procedure the use of information on those who were difficult to find. We differ from Behaghel et al. (2015) and Lee (2002) by developing a method that allows for non-monotonic differential attrition, using information from an intensive follow-up. The later relates to DiNardo, McCrary and Sanbonmatsu (2006)'s and Hull (2015)'s use of intensive tracking survey design features. Our approach extend the sample selection correction model based on observable characteristics by identifying, through the tracking protocol, a sub-sample of respondents similar to the sub-sample of attritors. That is, we assume that the ITT causal effect is homogeneous among individuals not found after regular tracking, after controlling for a large number of observables through the weights.

[5]See section G.1 in the Online appendix for a review of this literature.

This paper further builds on work studying attrition bias in household surveys in developing countries in non-RCT contexts. A number of studies use longitudinal datasets with low attrition rate to analyze differences between movers and stayers and to infer potential attrition bias on the analysis of the outcomes of interest (Thomas, Frankenberg and Smith (2001); Beegle, De Weerdt and Dercon (2011); Velasquez et al. (2010); Thomas et al. (2012)).[6] Overall, these studies agree on the fact that attritors differ from those who are found in observable characteristics. Alderman et al. (2001), Falaris (2003), Fuwa (2011) show that estimates are not necessarily biased even if attritors are different from stayers, but attrition bias can depend on the outcome of interest (Maluccio, 2004).[7]

To our knowledge, there is only one other paper specifically studying tracking protocols in the context of a RCT study in a developing country (Baird, Hamory and Miguel (2008)). Analyzing tracking in Kenya Life Panel Survey they compare ITT estimates of migrants that were tracked versus populations surveyed in their original locations and find evidence of heterogeneous treatment effects that are correlated with migration. Our paper starts from a similar finding, but then uses the information obtained from the intensive tracking phase to correct for attrition bias.

The next section presents the conceptual framework for our sample selection correction approach. Section 3 introduces the program, the evaluation design and the data collection used in the empirical application. Section 4 illustrates the sensitivity of the ITT estimates with and without inclusion of difficult-to-find respondents. Section 5 discusses the correlates of attrition and compliance to further understand the potential biases before introducing the new inverse probability weighting estimator. We then compare the results of the new estimator with other approaches, including standard IPW, bounds and proxy measures. Section 7 shows results applying the new IPW in three other datasets, and discusses generalizability. The last section concludes and discusses guidelines to evaluate

[6]For instance, Beegle, De Weerdt and Dercon (2011) show, with a household fixed effect model, that migrants moving out of their community of origin experienced 36 percentage points more of consumption growth than non-migrants household members between 1991 and 2004. They would have underestimated the growth in consumption by half of its true increase if they had focused only on individuals residing in their community of origin.

[7]Alderman et al. (2001) did not find any impact of attrition bias on anthropometric indicators in the Kwazulu-Natal Income Dynamics Study (KIDS), but Maluccio (2004) found evidence of attrition bias on expenditures using the same database.

cost trade-offs around intensive tracking.

2 Conceptual Framework

Consider a canonical, two-period ($t = 0, 1$), selection model in which the outcome variable y_{i1} is regressed on assignment to treatment (T_i),[8]

$$y_{i1} = \alpha + \beta T_i + \gamma x_{i0} + \epsilon_{i1} \tag{1}$$

where ϵ_{i1} is a mean-zero random variable, T_i is the treatment indicator, x_{i0} is a vector of individual and household characteristics observed for attritors and non-attritors at time 0 (at baseline).

Equation 2 specifies the process determining sample attrition or selection rule. It depends on the same independent variables (x_{i0}) as Equation 1 plus a vector of baseline variables (z_{i0}) affecting sample attrition but which are not part of the model of interest.

$$A^*_{i1} = \delta_0 + \delta_T T_i + \delta_1 x_{i0} + \delta_2 z_{i0} + \upsilon_{i1} \tag{2}$$

and,

$$A_{i1} = \begin{cases} 0 \text{ if } A^*_{i1} < 0 \\ 1 \text{ if } A^*_{i1} \geq 0 \end{cases}$$

The outcome y_{i1} is observed if $A_{i1} = 0$ and missing due to attrition otherwise. Then, the conditional mean of y_{i1} in the observed sample can be written as

$$\begin{aligned} E(y_{i1}|T_i, x_{i0}, z_{i0}, A_{i1} = 0) &= \alpha + \beta T_i + \gamma x_{i0} \\ &\quad + E(\epsilon_{i1}|T, x, z, \upsilon < -\delta_0 - \delta_T T_i - \delta_1 x_{i0} - \delta_2 z_{i0}) \end{aligned} \tag{3}$$

If there is correlation between both error terms, ϵ_{i1} and υ_{i1}, the last term in Equation 3 will be different from zero. Then, estimating Equation 1 ignoring Equation 2 will

[8] This section builds on Fitzgerald, Gottschalk and Moffitt (1998).

lead to biased estimates of β as the omitted variable $(E(\epsilon_{i1}|T,x,z,\upsilon < -\delta_0 - \delta_T T_i - \delta_1 x_{i0} - \delta_2 z_{i0}))$ is correlated with the variable of interest (T_i). Building on this logic, most often RCT studies focus on whether δ_T is equal or different from zero, that is whether attrition is balanced between treatment groups. If $\delta_T = 0$, the omitted term (the selection rule) is not correlated with T_i and β is an unbiased estimate of the treatment effect for the non-attrited population. That said, in case of treatment heterogeneity and when β covaries with υ_{i1}, estimates based on the non-attrited will not reveal the estimand for the entire target population (i.e. the external validity concern). Moreover, the conclusion on unbiasedness (internal validity) derives from the linear specification in Equation 2. If we allow, instead, that attrition can be differentially affected by treatment status depending on (x_{i0}) and (z_{i0}), the conclusion no longer holds.

This can be seen by expanding Equation 2 to allow for differential attrition in the selection equation,

$$A_{i1}^* = \delta_0 + \delta_T T_i + \delta_1 x_{i0} + \delta_2 z_{i0} + \delta_{T,1} T_i \times x_{i0} + \delta_{T,2} T_i \times z_{i0} + \upsilon_{i1} \tag{4}$$

$$\begin{aligned} E(y_{i1}|T_i, x_{i0}, z_{i0}, A_{i1} = 0) = \alpha + \beta T_i + \gamma x_{i0} & \\ + E(\epsilon_{i1}|T,x,z,\upsilon < -\delta_0 - \delta_T T_i - \delta_1 x_{i0} - & \\ \delta_2 z_{i0} - \delta_{T,1} T_i \times x_{i0} - \delta_{T,2} T_i \times z_{i0}) & \end{aligned} \tag{5}$$

In this scenario, and under the assumption of selection based on observables, unbiased estimates can, however, be obtained using weighted least square regression (Fitzgerald, Gottschalk and Moffitt, 1998; Wooldridge, 2002*a*).

To model the probability of sample selection on observables we make the following assumption on Equations 4-5,

Assumption 1.

1. *y_{i1} is observed whenever $A_{i1} = 0$*

2. A_{i1}, T_i, z_{i0} *and* x_{i0} *are always observed for all* i

3. υ_{i1} *is independent of* $\epsilon_{i1}|T_i, x_{i0}$

4. *For all* $z \in Z$, $x \in X$, $P(A_{i1} = 0|T_i, z_{i0}, x_{i0}) > 0$

Sample selection on observable characteristics implies that there is a vector of variables, z_i and x_i, that when interacted with T_i are strong enough predictors of attrition, such that the distribution of A_i given T_i, z_i, x_i and y_i does not depend on y_i, that is

Assumption 2. $P(A_{i1} = 0|y_{i1}, T_i, z_{i0}, x_{i0}) = P(A_{i1} = 0|T_i, z_{i0}, x_{i0})$

If the correlates of attrition are significantly different between treatment arms, that is $\delta_{T,1}$ and $\delta_{T,2}$ are different from zero, we obtain separate weights for each of the experimental group. The standard procedure to construct IPW consists of estimating the probability of being surveyed, conditional on a set of covariates (interacted with treatment), using the complete target population. Applying these weights adjust for the differences in baseline characteristics between treatment arms that arise because of attrition. Behrman, Parker and Todd (2009), for instance, construct IPW by treatment arm when estimating medium-term impacts of the PROGRESA/Oportunidades CCT program in Mexico.

This approach relies on the assumption of selection on observables for the entire target population. However, if those found during the intensive tracking phase are more similar to attritors, Assumptions 1 and 2 might be more plausible if selection is modeled for the intensive tracking only. Thus we relax these assumptions and model selection among those individuals that were hard to find. We start from the insight that individuals tracked during the intensive phase (whether found or not) are similar on observed characteristics. We define overall attrition as $A_{i1} = A_{i1}^{ITP} + (1 - A_{i1}^{RTP})A_{i1}^{ITP}$, where A_{i1}^{ITP} is an indicator variable taking the value of 1 if individual i was not found after conducting intensive tracking, and A_{i1}^{RTP} is an indicator variable taking the value of 1 if individual i was not found after conducting regular tracking. We can relax Assumptions 1.4, and 2, and assume:

Assumption 1.4'. For all z $\in Z$, x $\in X$, $P(A_{i1}^{ITP} = 0|T_i, z_{i0}, x_{i0}, A_{i1}^{RTP} = 1) > 0$

Assumption 2'. $P(A_{i1}^{ITP} = 0|y_{i1}, T_i, z_{i0}, x_{i0}, A_{i1}^{RTP} = 1) = P(A_{i1}^{ITP} = 0|T_i, z_{i0}, x_{i0}, A_{i1}^{RTP} = 1)$

The basic assumption underlying our estimation strategy is that individuals that were hard to find are similar in unobservables to those that ultimately were not found. We modify the IPW estimates by exploiting these similarities, and construct weights using only the individuals tracked during the intensive tracking phase.

We predict the probability to be found using only those individuals who were missing after regular tracking,

$$\hat{p}_{i1}^{ITP} = P(T_i, z_{i0}, x_{i0}, \hat{\delta}_T^{ITP}, \hat{\delta}_1^{ITP}, \hat{\delta}_2^{ITP}, \hat{\delta}_{T1}^{ITP}, \hat{\delta}_{T2}^{ITP}) \tag{6}$$

Respondents interviewed during the RTP are assigned a $P(A_{i1} = 0|T_i, z_{i0}, x_{i0}) = 1$. We obtain the target population density function by weighting the conditional density in the ITP sample using $w_{i1}^{IT}(T, z, x) = \left[\frac{1}{\hat{p}_{i1}^{ITP}}\right]$ and weighting the conditional density in the RTP sample using $w_{i0}^{RT}(T, z, x) = 1$. Compared to the standard IPW, the new IPW accounts for differential sample selection during different tracking phases. We predict the probability to attrit from a sample of attritors and respondents who were also difficult to track (those found in the ITP) and hence focus on characteristics that differentiate both sub-groups. Given that the RTP sample is larger than the sample of ITP, the "differential" characteristics between the ITP sample and the attritors sample are diluted when we apply the standard IPW.

The sample size of the intensive tracking sample will, of course, affect the predictions in the intensive tracking sample and hence the weights to be used. Larger samples will allow to account for more potential drivers of selective attrition, without running into overfitting concerns in the prediction model. Depending on the total number of hard-to-find individuals, this could be one reason to intensively track all of them, rather than a random subset, as is commonly done.[9] On the other hand, when sample sizes are large, or potential drivers of selective attrition are limited, tracking only a random subset allows

[9]Intensively tracking only a random subset also reduces statistical power of the impact estimates (as compared to full tracking), and more so in case of individual randomization or cluster randomized trials with low intra-cluster correlations.

to reduce costs and still obtain reliable predictions and weights.

3 Red de Protección Social: Program design, Evaluation and Data[10]

3.1 Program design and Evaluation

The Red de Protección Social (RPS) was a conditional cash transfer program launched in 2000 targeting households living in rural poor Nicaragua. The design of the program closely resembles the well-known PROGRESA/Oportunidades program in Mexico and consisted of cash payments to the main female caregiver in the household of approximately 18 percent of total annual household expenditures. Transfers were conditional, and households were monitored to ensure that children were attending school and making visits to preventive health-care providers.

To conduct a rigorously randomized evaluation of the program, 42 localities from 6 municipalities were randomized into treatment and control groups at a public lottery (stratified by poverty level). The program started in the 21 treatment localities in mid 2000 and lasted for 3 years (hereafter, early treatment localities). In 2003, the experimental treatment localities stopped receiving the transfers, while the program started in the experimental control localities (which hence became the late treatment localities). This group received transfer during the following three years. All households received sizable "food" transfers, a fixed amount independent of the number and age of family members. Households with children between 7 to 13 years old who had not finished the first 4 grades of primary school got an extra education transfer conditional on school attendance.

We exploit the experimental design and the long-term follow up data to study how attrition affects the impact estimates for boys 9-12 years at baseline, following the identification strategy in Barham, Macours and Maluccio (2017). This cohort had greater program exposure in the early treatment localities than in the late treatment localities

[10] See Flores and Maluccio (2005) for additional details on the program and the experimental design.

due to the eligibility criteria for the education transfer and pre-program school dropout patterns. It includes children that were young enough to be eligible for the education transfer if they were living in a early treatment localities in 2000, but too old to receive the education transfer when the program phased-in to the late treatment localities in 2003. Barham, Macours and Maluccio (2017) use the experimental variation in timing to estimate the long-term differential impacts of the program on a wide set of education and labor market outcomes. In this paper we investigate the implications of attrition for estimates of grades attained and participation in off-farm employment, two of the main outcome variables of the long-term evaluation.[11]

3.2 Survey Data

We use data from a census conducted before the program started in May 2000 and a follow-up survey conducted in 2010. The follow-up survey targeted the 1,756 households randomly selected for the short-term evaluation of the program, as well as a sample of 1,008 households drawn from the baseline census in the early and late treatment localities and added in 2010 to increase the sample size for certain age groups. These groups were over-sampled to maximize the difference in the potential length of exposure to the program at critical ages between the early and late treatment groups. The new sample was randomly selected using the census data from 2000.[12] The 2010 sample includes all households that contain the original beneficiary of the program. In addition, if an original panel household member under 22 (in 2010) had moved out of the household by 2010, their new household (the split-off household) was added to the sample. During the follow-up the survey team interviewed 2,505 original households and 1,375 new households, including both local and long-distance migrants. Substantial effort was made to track individuals to limit attrition due to migration and household split-off. Households and individuals in the target group were tracked across Nicaragua and to Costa Rica. Multiple visits to the original localities reduced attrition due to seasonal migration.

[11] Off-farm employment is measured as a dichotomous variable that takes value one if the individual is economically active (in wage or self-employment) outside of the family farm, and zero otherwise.

[12] To keep the sample representative of the target population, all estimates include sample weights constructed at the locality level.

3.3 Tracking and Tracking Costs

The tracking process lasted almost 2 years. During the first phase of data collection, from November 2009 to March 2010, all sample individuals were tracked in their localities of origin and some migrants were followed to other localities within the 6 municipalities. We refer to this phase as *Regular Tracking Phase* (RTP) as it is similar to the most used tracking protocol in longitudinal surveys, even if it already includes information on some migrants. In April 2010 the second phase was launched and non-found target individuals were tracked intensively, to other regions or to Costa Rica. During this phase, *Intensive Tracking Phase* (ITP), the enumerators also went back to the localities of origin for regular updates on the destination information and to survey returned temporal migrants (see Appendix C for more detail on the tracking protocol).

The RPS baseline population census included questions about the characteristics and composition of the household, education and economic activities of household members, ownership of durable goods, land property and information on agriculture activity. The questionnaire in 2010 includes sections on education and economic activities for all household members, as well as a large section on permanent migration including information about where and how to locate migrants. It also included a limited set of questions on the education and occupation of all baseline members who had permanently moved out, asked to the household head or the main program beneficiary (hence typically the father or mother of the absent individual). We exclude this proxy information on permanent migrants in most of the analysis, but return to it in Section 6.

To evaluate the cost of the tracking process we calculate the number of enumerator workings days (that is the number of days the team worked times the number of enumerators in the team at each moment). During the regular tracking phase, the team worked 91 days, and the cost to find and interviewed the RTP sample was 1,486 enumerator days. Note, that this number also accounts for the cost of gathering information on migrants' destination. To track and interview the ITP sample the enumerator team worked 218 days and the total number of enumerator days on this phase reached 905 (see Figures

1-2).[13]

3.4 Survey Attrition

Final attrition rates for males in the cohort of interest are low for a 10 year panel despite high mobility. Around 40 percent of the target sample had permanently moved to another location between the baseline survey and the follow up survey in 2010. Another 24 percent temporarily migrated for work or study at least part of the last 12 months. After intensive tracking the final attrition rate for the targeted sample is 10.19 percent, corresponding to an annual rate of 1 percent (Table 1).[14]

The top panel of Table 2 shows response rates by treatment group at different stages of the tracking process. Response rates are not significantly different between the early and late treatment groups at the different stages.[15] After RTP attrition rates were still relatively large, 26 percent, but differences between treatment arms are not significantly different from zero. The third row of Table 2 shows the response rates after conducting ITP conditional of not being found during the RTP. Around 60 percent of those not surveyed after conducting RTP were found during the ITP, with the differences between treatment arms again not significantly different from zero.

The bottom panel of Table 2 shows that response rates after RTP no longer appear balanced once we consider observable subgroups (based on baseline characteristics). Hence the overall balanced rate obscures that the type of individuals that are attriting are quite different between the early and late treatment groups, leading to analysis samples that are no longer comparable.

Table A1 in Appendix A shows how attrition affected balance of variables observed in the baseline census. The table shows that the randomization resulted in very few

[13] Accounting for survey breaks, the regular tracking phase spanned a 5-month period, and the intensive tracking phase spanned 1.5 years. As Figure 1 and 2 show, most of the observations in the intensive tracking phase where collected in the 5 months directly following the regular tracking phase. This difference in timing between the regular and most of the observations of the intensive tracking phases is small compared to the 10-year period since baseline, and is unlikely to be driving the results.

[14] Attrition includes those who have migrated and those who refused to be interviewed, which account for less than 0.1 percent of those non-respondents.

[15] That said, the power calculations underlying the randomized design were not done to be able to detect selection into attrition and hence our study, as almost all other studies, is underpowered to capture such differences in response rates.

significant differences between the early and late treatment group on the full sample (column 1), as expected. After regular tracking, however, a number of additional baseline variables were off balance, in particular related to parental education and household demographics (column 3), in line with results in Table 2. This suggests that only regular tracking would have introduced potential important selection bias. Notably, column 2 shows that after intensive tracking, these imbalances are no longer there, and the only remaining variables that are significant are the few that are significant for the full baseline sample (column 1). This is consistent with boys found during the intensive tracking phase being different in observed characteristics from boys found during the regular tracking phase.

4 Intent-to-Treat Estimates: Education and Off-farm Employment

We next show the ITT estimates of the differential impact of RPS on education and off-farm employment, comparing estimates obtained after regular tracking with those obtained after complete tracking. The former represents the results that would have been obtained if only common tracking rules would have been applied to the survey sample. The later represents the benchmark estimate after exhaustive tracking but without further correction for remaining attrition. We also separately show ITT estimates for the subsample tracked during the intensive phase.

Equation 1 takes the following form:

$$Y_{i2010} = \alpha + \beta T_i + \gamma X_{i2000} + \epsilon_i \tag{7}$$

where Y_{2010} is the outcome of interest in 2010, T is an ITT indicator that takes value of one for children in localities randomly assigned to early treatment and zero otherwise, and X_{2000} is a set of controls at baseline.

Table 3 shows the ITT estimates for boys ages 9-12 for samples completed at different stages during the tracking process on the grades of education attained (top panel) and on off-farm employment (bottom panel).[16] The table shows results under two specifications for X: the 1st model includes strata, three monthly age fixed effects and baseline education[17] (columns (1), (3) and (5)); the 2nd model adds controls for distance to school, number of children 0-8 and 9-12 in the household, estimated per capita consumption and estimated per capita consumption squared, regional fixed effects as well as a vector of covariates that was off-balance after the relevant tracking phase - see note under Table 3 (columns (2), (4) and (6)).[18] Estimates on the full sample show that boys coming from localities randomly assigned to early treatment have 0.427 more grades attained than boys from the late treatment group. Including the variables that were not balanced at baseline and other baseline controls correlated with the outcome and regional fixed effects decreases the point estimate to 0.319. Overall, the results show that the ITT point estimate after CTP is relatively sensitive to the inclusion of baseline controls, despite the randomization.

The next two columns show that the size and sign of the estimate is driven by those found during the RTP. If the follow-up had been completed after regular tracking, the point estimates would have been larger, reaching 0.613 in the first specification, significantly different from zero at the 1 percent level. The baseline balance test after RTP showed that both groups differed in many dimensions, due to selective attrition. Adjusting for these imbalances and other variables decreases the point estimate to 0.350. The last three rows report p-values for testing the equality of coefficients at different stages during the tracking process. The ITT estimate after RTP is 44 percent larger, and significantly different from the final estimate using the CTP sample, in the first specification. Differences between CTP and RTP are smaller and not significant with the expanded set

[16]The sample includes 1,006 individuals found and with information on grades attained and off-farm employment in 2010. The sample does not include 15 deceased individuals.

[17]A set of dummies indicating whether the individual had 1, 2, 3 or at least 4 years of education at baseline.

[18]Table A2 shows additional specifications to present results for the common approaches followed in the RCT literature, going from only controlling for stratification to including more information in the regression model (Athey and Imbens, 2017; Deaton and Cartwright, 2016). Following Athey and Imbens (2017) we also re-run the analysis using a transformation of the continuous covariates into indicator variables. To do so, we replace the categorical and continuous covariates with a set of binary variables indicating whether individuals is above the median for each of those variables. Results are generally robust, see Appendix E.

of controls.

The last two columns show the ITT estimates for the sample of individuals found during the ITP. None of the point estimates are significantly different from zero and, if anything, the sign of the coefficients suggests that those in the early treatment group end up with slightly less grades of education attained. These estimates should clearly not be interpreted as causal, as selection into this sample is different for the two treatment groups, but they help explain why the ITT estimates on the complete sample are smaller than after RTP.

A broadly similar pattern emerges for off-farm employment. The results show that boys assigned to early treatment are about 6 percentage points more likely to be off-farm employed relative to boys in the late treatment group. Among boys found during the regular phase of the tracking protocol, ITT estimates are about 28 percentage points larger and these differences are significant at the 1 percent level even for the specification with full controls. Among boys found during the ITP ITT estimates are negative, indicating that those in the late treatment group are more likely to have an off-farm job.

Hence selection at different stages of the tracking process affects ITT estimates for both outcomes in the same direction and with similar order of magnitude. Not including those found during the ITP leads to a substantial overestimate of the basic ITT effects in the basic specification. Including baseline controls reduces the difference for grades attained but not for off-farm employment. Intensive tracking of course comes with a cost, which may need to be weighted against the benefits of reducing attrition bias on the ITT estimates. To assess the cost in terms of days and number of enumerators, Figures 1 and 2 shows the evolution of the ITT point estimates on each of the outcomes of interest (vertical axis) as a function of the number of enumerator days during the intensive tracking phase.[19,20] The figures show that individuals found during intensive tracking in close-by regions and Managua are driving point estimates down. They also

[19]An enumerator day is defined as any working day in which the team of enumerators worked after RTP (March, 23th 2010) times the number of enumerators in the team at each date.

[20]These numbers do not account for the field work done during the RTP that also includes collecting information on migrants destination but it gives a lower bound estimate of the cost in terms of field work. Costs are calculated for the entire sample of 6,000 individuals that were tracked, of which 299 are boys 9-12 at baseline.

show that estimates stabilized in the later part of the intensive tracking process. While it probably would have been hard to predict this particular pattern prior to the intensive tracking phase, the graphs are consistent with heterogeneous treatment effects. The sensitivity of the estimates to the additional controls further may point to bias due to selective attrition.

In the following sections we address this potential remaining attrition bias.

5 Inverse Probability Weights

5.1 Correlates of attrition

Results in Section 4 confirm the importance of understanding the correlates of attrition to make informed assumptions about the nature of selection into the final sample, even when response rates are balanced by treatment group. To do so, we consider both the context and households' reaction to the program. Program participation can induce different types of individuals to migrate and attrit in early and late treatment, even if on average the same number of people leave the sample. Table 2 illustrated several examples of such non-monotone differential attrition based on observables subsamples. The probability to find any particular individual is affected by various prior decisions by that person and his household. Individuals that have moved out of the study region, before, during or after the program will be harder to find, as are individuals who temporarily migrate for work or family reasons. These migration decisions can be affected either directly or indirectly by the randomized exposure to the program studied, but could also capture the heterogeneity of the population.

It seems plausible that the intervention studied affected migration positively for some individuals, and negatively for others, and this heterogeneity is likely to affect the impact estimates. The CCT program had the specific objective to increase educational attainment for the target population, and transfers were conditional on the presence and attendance of the boys to school. The transfer package in general, and the conditionalities in particular, a priori should have reduced migration during the program years. On the other hand, to the extent that the program effectively increased educational attainment, this could have increased or decreased migration after the end of the program.[21]

While the program analyzed, and the differential timing of transfers in the early and late treatment villages are specific to our study, the potential of large interventions, such as CCTs, to affect migration behavior, and hence the probability to attrit both positively and negatively, is much more general. Similarly, covariates of attrition likely differ from context to context, but broadly speaking, differences in socioeconomic status (SES), existing networks, family structures, and temporary residence are likely candidates to help explain differential attrition in many contexts. Appendix B discusses the logic underlying the choice of covariates for the program studied.

Table 4 shows average baseline values of household and individual characteristics. It compares individuals found after the CTP, after the RTP, for the subsample of those found during the ITP and finally for those missing. The last two columns show differences in means between the CTP sample and ITP sample with respect to the attritors sample. The differences are significantly different from zero for a number of observable characteristics, confirming that they are likely correlates of the migration decision or of the accuracy of the migrant destination information obtained, in line with the reasons for migration and attrition discussed in Appendix B.

Table 4 further shows that those found in the intensive phase are much more like those not found along many of these dimensions. For most of the indicators of socio-economic status for which there are significant differences between the found and the attrited sample, differences are smaller and often not significant anymore when comparing the attrited with those found in the intensive phase. For instance, those found during the intensive phase are similar in remoteness, productive assets, and land ownership to the attritors, which is not the case for those found in the regular phase. Along the same lines, all proxies of family networks and temporary residence show that those tracked in the intensive phase share more commonalities with those ultimately not found, than those

[21] Migration may have increased because of increased job opportunities outside of the villages of origin, or even because individuals migrated to continue their education elsewhere. Yet, if increased education increases the returns to self-employment activities in the program villages, the effect could also be the reverse. This could be more relevant for households with complementary productive assets.

found during the regular phase.

The similarities between those intensively tracked and found versus not-found are not surprising, as almost all of these individuals took the decision to move out of the villages of origin, and often moved to locations in different municipalities and departments (see Table 5). Compliance rates among those tracked in the intensive phase are lower than those found during the regular tracking, consistent with migration that occurred before the program started or during its initial phases. Compliance is indeed much lower in areas with very early attrition, and for temporary residents, while it is notably higher for households with large family networks. Compliance more generally is correlated to many of the same baseline observables as attrition, indicating that attrition weights may in part capture the treatment heterogeneity related to different levels of compliance (see discussion in Appendix B).[22]

Overall Tables 4 and 5 show that attritors and individuals surveyed during the intensive tracking are more similar to each other than attritors and RTP individuals. Similar findings have been reported for other panel surveys with intensive tracking protocols.[23]

The differences between the attritors and non-attritors, and the similarities between those found or not found in the intensive phase, are important as program impacts may well differ along many of these same dimensions. Decisions on investments in children's education or work could be different in households with different SES background, in non-nuclear households, or for temporary residents. Moreover it is not a priori obvious how this would affect program impacts, as the CCT may have induced children to get more education when they otherwise would not have, or, on the contrary, may re-enforce existing differences. Differences in compliance also directly affect ITT estimates.

The correlates of attrition and of compliance further differ between the two experimental groups (see Appendix B). This is particularly the case after regular tracking. Estimates

[22] Note that ToT estimates in this context would be hard to interpret as both treatment groups eventually receive the intervention, and in each group some households decide not to take it up or comply with conditions. Non-compliance with their treatment assignment hence is different than typical non-compliance in RCTs with a pure control group.

[23] Using data from IFLS, Thomas et al. (2012) show that longer distance migrants have more in common with those not found in the follow-up than those who didn't move. Beegle, De Weerdt and Dercon (2011) find similar patterns in their analysis on economic mobility in Tanzania.

after regular tracking hence likely would be biased, consistent with the large difference in ITT estimates after RTP and CTP shown in Table 3. Differences become smaller after intensive tracking, but some remain. Furthermore, attrition remains selective, as many baseline characteristics remain significant predictors of attrition. This suggests that even the ITT estimates after complete tracking likely do not reflect intent-to-treat estimates for the entire target population in the presence of heterogeneous treatment effects. We therefore calculate weights to correct for selective attrition.

5.2 Obtaining Probability Weights using information from tracking

The intensive tracking strategy provides valuable information to calculate weights for the attrition selection correction described in Section 2. We will compare the new IPW estimates with regular IPW estimates obtained both only considering observations found during regular tracking, and all observations.

For these different estimates, and following Fitzgerald, Gottschalk and Moffitt (1998) we account for a set of baseline variables that may be driving selection (henceforth, Z)[24] and a vector of baseline covariates that form part of a basic model of education (henceforth, X). While the categorization of the variables is somewhat arbitrary, we think of the SES variables (such as strata fixed effects and household assets) as mostly capturing X, the variables related to the temporary nature of residence, networks and regions as Z, while the demographic variables can be classified as either X or Z. This gives as potential predictors a wide set of baseline variables capturing the socio-economic status, the demographics, the baseline networks, and the possible temporary nature of the households baseline residence (see Table 4).

As there is a wide set of observed characteristics to consider, and as there are relatively few observations not found after intensive tracking, we follow Doyle et al. (2016) to reduce the set of predictors.[25] We first estimate bivariate regressions in which each potential

[24]See Appendix B for a discussion on the theoretical underpinnings and contextual factors for the selection of variables included in Z.

[25]An alternative set of predictions for weights is obtained using LASSO, which yields broadly similar

predictor was tested to determine whether a significant difference existed between those found and those not found. All estimates use population weighted observations and standard errors clustered at the locality level. The testing was conducted separately for the early and late treatment group.[26] Results can be found in Tables B1, B3 and B4 in Appendix B. Any measure found to be statistically significant for the early or late treatment group was retained as a potential predictor. We then estimate the probability of being found on this set of baseline predictor variables separately for each experimental group. In order to account for collinearity between measures, the baseline predictor set was further restricted by conducting stepwise selection of variables with backward elimination and using the adjusted R-squared as information criteria. The strata and regional fixed effects, as well as 6-monthly age dummies were included as fixed predictors in all regressions.

In the last step, we estimate the probability of being found for both early and late treatment groups together, keeping only the predictors as indicated by the stepwise procedure. Following Thomas et al. (2012), we also included interviewers characteristics (fixed effects for the team that first visited a village during the regular tracking) to capture differences between teams in effectiveness of obtaining information for tracking. Table B5 in Appendix B shows the results, which confirm that selection is not random in any of the tracking phases. The estimate for the intensive tracking phase in particular has good predictive power (the linear probability model has an R-squared of 51 percent), and the predictive power is much higher than for estimates on the full sample or the regular tracking sample (with the adjusted R-squared being at least two times higher). Hence restricting the sample for estimating weights to the ITP, increases the predictive power of the model and reduces measurement error in the estimates of the weights. Selection after CTP is driven by both X and Z variables, though Z variables appear strongest (i.e: variables related to location, household property and locality ex-ante attrition rates).[27]

results. See Appendix F for discussion and results, and Imbens (2015) for a related approach.

[26] We include estimates of early and late treatment localities in the same regression but interact the variable of interest with the treatment status to obtain separate estimates for early and late treatment group.

[27] See the bottom panel of Table B5.

Meanwhile, selection after RTP is mainly driven by the X variables, though demographic, and Z variables also play a role. In contrast and in line with the findings for the CTP sample, among those tracked during ITP, selection between those found and those not found is mainly driven by Z variables, with a less strong role of X or demographic variables. The results further suggest that selection differs by treatment group during each of the tracking phases. This finding points to a potential threat to the monotonicity assumption needed for other attrition selection corrections (Lee, 2009; Behaghel et al., 2015).

5.3 Results with Inverse Probability Weights

Table 6 shows WLS estimates for assignment to early treatment on grades attained (top panel) and on the probability of off-farm employment (bottom panel). The first two columns show WLS estimates with weights capturing selection during the intensive tracking phase, i.e. the new IPW estimates. Final weights for the new IPW vary between 1 and 35.[28] We compare these estimates with the OLS estimates in the first 2 columns of Table 3. We also compare with WLS estimates using weights estimated with the entire sample (columns 3 and 4). In the bottom of each panel, we show WLS estimates only using the sample from regular tracking.[29]

The ITT estimates on years of education using the new IPW, which corrects for selection on observables during intensive tracking, are smaller than the un-weighted ITT estimates in the first specification. In the models in which we only control for strata, age, and education at baseline, the new WLS is 0.363. Adding the full set of baseline controls (column 2) reduces the point estimate of the new IPW slightly to 0.317, significant at the 10 percent level. In contrast, applying standard IPW to the CTP sample leads to much larger differences in the point estimates in the two specifications (columns 3 and 4). The WLS estimates appear as sensitive as the OLS estimates to the inclusion of different set

[28] Results are robust to dropping of observations with the highest weights. Figure A1 in the appendix show the distribution of the standard weights and the new weights for the complete sample.

[29] Wooldridge (2002*b*) shows that computing the asymptotic variance of WLS estimates ignoring that the probability to be surveyed was predicted in a first-step leads to larger standard errors. Our estimated standard errors should therefore be conservative.

of controls, suggesting that the standard weights by themselves may not correct for all the selective attrition.

Comparing the new IPW (0.363) with the ITT estimates (0.427) after complete tracking, shows that the proposed attrition correction reduces the point estimate for grades attained. This is not the case when comparing the new IPW estimate and ITT estimates with the full set of controls, suggesting that the control variables in this sample - and once intensive tracking has been done - allowed to correct by themselves for the remaining selection. The difference in the first specification is of interest, however, as it corresponds to the more standard ITT estimates used for RCTs, with limited and discrete controls in the first column, in line with Athey and Imbens (2017) recommendation.

The education estimates after the regular tracking, are even more sensitive to changes in the set of control variables, further consistent with the controls potentially helping to address bias due to selective attrition. In contrast, the results using the new IPW appear the most robust to different specification of the controls, suggesting that the improved weights already take care of the sample selection correction.[30] Overall, the attrition correction from adding the controls and estimating the new IPW is substantial. That said, the biggest difference in estimates is obtained when moving from the regular tracking to the full tracking sample, suggesting that putting effort in tracking remains the first best response to limiting attrition bias.

In the case of off-farm employment, all estimates after CTP appear relatively robust to the inclusion of different covariates. However, when we compare the estimates after CTP to the estimates after RTP we find that applying the standard inverse probability weights to the sample after RTP and including baseline controls in the specification leads to higher point estimates. The adjustment is not as large as for grades attained, which could be explained by more homogenous treatment effects for this outcome. Nevertheless, the results for off-farm employment confirm the earlier finding that estimates after RTP appear to overestimate the treatment effects. The lower compliance among individuals

[30]Table A3 in Appendix A shows the results for the four specifications in Table A2. Results in Appendix E show that point estimates for the new IPW are generally robust to include discrete covariates instead of continuous ones, while the point estimates for the standard IPW on complete or regular sample are more sensitive.

that were hard to find offers an intuitive possible interpretation for this finding. Overall, the result illustrates that having a balanced attrition rate between treatment groups does not guarantee unbiased ITT estimates valid for the entire target population, and that intensive tracking is potentially important to limit such bias.

6 Other methods to deal with attrition

We now compare the estimates with other methods to correct for attrition bias. First, we compute bounds of the ITT estimate under different assumptions on the distribution of attritors. Second, we exploit reported information by non-migrant household members on education outcomes of migrants and test whether using proxy information would have resulted in similar ITT estimates. We focus on the specification with strata, age and baseline education controls only.[31]

6.1 *Ex-post:* Non-Parametric Bounds

We calculate three types of bounds, each reflecting different assumptions about the distribution of the treatment effects. First, we follow Horowitz and Manski (2000) and construct bounds assuming a worst-case scenario (1st case). To compute lower (*upper*) bounds we impute minimum (*maximum*) value of the outcome in the non-attrited distribution to those not found from early treatment localities and the maximum (*minimum*) value of non-attrited distribution to the attrited from late treatment localities. For the second group of bounds, we follow Kling, Liebman and Katz (2007) and impute for the lower (*upper*) bound the mean minus (*plus*) 0.75 standard deviations of the non-attrited early treatment distribution to the attritors from early treatment localities and the mean plus (*minus*) 0.75 standard deviations of the non-attrited late treatment distribution to the attritors from late treatment localities. We repeat the same exercise using 0.50 and 0.25 standard deviations, to reflect other common values found in the literature. Third, assuming monotonicity in the selection rule, Lee (2009) proposes to construct tighter

[31]Hence estimates are to be compared with the 1st specification in Table 3 and 6. Estimates for the other specifications can be found in appendix Tables A4-A5.

bounds by trimming the sample of respondents such that the share of observed individuals is equal for both groups, early and late treatment.

Table 7 shows lower and upper bounds for grades attained (top panel) and off-farm employment (bottom panel). The intervals between the Manski (worst-best) bounds after both RTP and CTP are large and uninformative for both outcomes. Intervals imposing more restrictive assumptions on the distribution of treatment effects (following Kling, Liebman and Katz (2007)) are narrower, although the ranges between the upper and lower bound are still large for the less strict bounds. And applying the most restricted bounds (± 0.25 $s.d$) after RTP leads to a lower bound that is higher than the new IPW estimate. Following a logic similar to the new IPW estimates, we also estimated bounds using only the observations tracked during the intensive phase, and then applying those bounds to the estimates of the full sample. Those bounds are a bit tighter but still largely uninformative.

Lee bounds after RTP are tighter than worst-case scenario bounds. Conducting intensive tracking reduces the percentage of trimmed observations by one fourth and narrows the bounds further. But an inconvenience of the Lee bounds in small samples is that it is hard to account for a large set of control variables. This could potentially explain why Lee bounds are above the OLS estimates on grades attained. More generally, the Lee bounds are above both the new IPW estimates and the OLS estimates with controls. To interpret this, we return to the monotonicity assumption underlying the Lee bounds. Bounds around the treatment effect are useful when the sample does not suffer selection based on observable characteristics, other than those accounted for in the cells for bounds calculations. Section 3.4 suggests, however, that in our sample we do have selection on a relatively large number of observable characteristics and that this selection is not homogeneous at different stages during the tracking process, hence raising doubts on the validity of the monotonicity assumption.

6.2 *Ex-ante:* Proxy information

An alternative approach to correct for attrition "ex-ante" is to collect proxy information on household members who no longer belong to the baseline household. Following Rosenzweig (2003), we compare self-report (y_{i2010}) and proxy reports (y_{i2010}^{proxy}) to analyze data reliability.

Hence, we estimate,

$$y'_{i2010} = \alpha + \beta T_i + \gamma x_{i2000} + \epsilon_{i1} \tag{8}$$

where,

$$y'_{i1} = \begin{cases} y_{i2010} \text{ if } A^*_{i2010} < 0 \\ y_{i2010}^{proxy} \text{ if } A^*_{i2010} \geq 0 \end{cases}$$

As basic information on education and economic activities was collected for all baseline household members, independently on whether they still lived in the household in 2010 or not, we have double information on migrants who were subsequently interviewed in their new household. The main source of information corresponds to the data collected in the household where the individual is currently living, while the "proxy information" is information reported by a member from the household where the individual was living in 2000. We analyze reliability separately for different tracking phases, as distance and time could affect the accuracy of the information.

We have double information on 18 percent of the final sample and 42 percent of those found in the intensive tracking. Differences in the availability of proxy information and on its level of accuracy between experimental groups are small and not significant (see Tables A6 and A7 in Appendix A). For grades attained, the percentage of correct proxy information is higher among those who were surveyed during the RTP (63 percent) than among those surveyed during the ITP (51 percent), consistent with these respondents being closer and better connected to their households of origin. The size of the bias (proxy minus self-report) is small and negative, in both RTP and ITP sub-samples, indicating that if anything households tend to underestimate the level of education of migrants. In contrast, proxy information reported on off-farm employment is more accurate for those found during ITP (72 percent) than for those found during RTP (60 percent). In both cases households tend to underreport off-farm employment, but the size of the bias is larger in the RTP sub-sample.

To analyze the implications of using proxy information, we use proxy information for boys not found at the end of the regular or intensive phase together with self-reported information. First, we consider those from whom we have double information. The first two columns of Table 8 show the benchmark ITT estimates for the sample of respondents after CTP and RTP on grades attained (top panel) and on off-farm employment (bottom panel). Columns 3 and 4 report ITT estimates on the RTP sample plus the sample of respondents found during the ITP from whom we have double information (42 percent of the sample of ITP respondents). Column 3 shows results with self-reports (i.e. estimation of Equation 1) and column 4 shows results with proxy reports (i.e. estimation of Equation 8). The results show that for the ITP sample of respondents' proxy information would have led to smaller point estimates. Comparing these results to the estimates after RTP we observe large differences, as adding the sample of respondents from the ITP with double information reduces the ITT estimates on grades attained by about 0.1.[32] These results support previous findings suggesting that RTP estimates overestimate the value of the ITT estimates on education. The bottom panel shows that adding proxy information for off-farm employment on respondents found during the ITP raises the ITT estimate after RTP slightly, while using self-reported information on the ITP respondents with double information does not affect the ITT estimate after RTP.

Columns 5 to 6 extend the sample to include all the attritors from whom we have proxy information reported by non-migrating household members. Column 5 reports the ITT estimates obtained by adding the sample of attritors from whom we have proxy information to the CTP sample and in column 6 we use proxy information on the sample of attritors for ITP respondents from whom we have proxy information. This last exercise gives us the estimates that we would have found if the ITP would not have been conducted. Using proxy information on education and off-farm employment for the sample of

[32]Estimates with full set of baseline controls are qualitative similar, but differences are smaller.

attritors to estimate Equation 8 gives point estimates similar to the ITT treatment effects estimated using the sample of respondents (after CT) (column 5). But the last column shows that we would have overestimated the ITT for both grades attained and off-farm employment if using only proxy information on the entire ITP sample. The estimates of grades attained is 9 percentage points higher than the benchmark ITT estimate (but smaller than, the ITT estimate after RTP). And the estimate for off-farm employment is 50 percent higher than the benchmark ITT.

These results suggest that proxy information on attritors can help correct for some attrition bias. That said, proxy information on those not found during RTP would still have led to an overestimate of the ITT estimates in this study, suggesting the approach also has its limitations. Analyzing proxy information on those that eventually are found during intensive tracking can help sign the potential bias introduced by the proxy report.

7 External validity

To illustrate the use of the new IPW approach proposed in this paper, we apply the methodology to three other RCT-panel datasets.[33] For each dataset, we follow the steps lined out in Section 5.2 to calculate the new IPW, keeping the error structure, control variables, and population weights the same as in the original papers. We show estimates for the ITT, the new IPW, and the standard IPW to demonstrate how the attrition correction affects the estimates.

Macours, Schady and Vakis (2012) estimate the impact of a one-year cash transfer program ("Atención a Crisis") on early childhood development. The program was implemented in a different region of Nicaragua than RPS, with 56 localities randomly selected to receive the program and 50 localities randomly selected as control. Baseline data was collected in April-May 2005 and endline data was collected between August 2008 and May

[33] Among the RCTs included in the literature review (see Appendix G.2), five datasets satisfy the basic requirements to apply our methodology, that is: (i) the endline included an intensive tracking phase and information on the subsample that was tracked intensively; (ii) the baseline data includes a vector of variables with enough information to estimate a model of attrition. For Macours, Schady and Vakis (2012), Blattman, Fiala and Martinez (2014) and Blattman, Fiala and Martinez (2020) we had access to the relevant data.

2009. We focus on the impact of the conditional cash transfer on the primary outcome in Macours, Schady and Vakis (2012), the overall index of cognitive and socio-emotional outcomes, and consider the sample of children born at baseline and under 5 years old. The response rate after CTP was 0.952 in both treatment and control localities. Attrition rates after RTP were substantial, 29 percent in treatment and 33 percent in control, with the difference significant at the 10 percent. Attrition in the regular tracking phase was due to non-response or inability to locate the children at the moment of the survey visit, in addition to migration. During the ITP 85 percent of those not found during the regular tracking phase were found and interviewed.

Table 9 shows how the new IPW estimate (2nd column) compare with the ITT (1st column) and a standard IPW estimate (3rd column). The top panel shows results for the complete sample and the bottom panel shows results for the sample after conducting regular tracking. Results show that ITT estimates declined as a result of intensive tracking, but that they are similar for the ITT and the IPW estimates. For the results with complete tracking, this may not be surprising since with less than 5% of observations missing, selection is minimal. Moreover, given the young ages of the children and the outcomes studied, it seems plausible there isn't necessarily much correlation between the remaining migration and program impacts. The similarity between the ITT and the IPW after regular tracking does imply that the regular IPW did not manage to introduce the relevant sample correction.

The other two datasets were collected for the mid-term and long-term evaluation of Uganda's Youth Opportunities Program (Blattman, Fiala and Martinez (2014) and Blattman, Fiala and Martinez (2020)), through an RCT that randomly allocated 265 youth groups to treatment and 270 to control, stratified by district. We focus on the impact of the program on income in 2012 and 2017, 4 and 9 years after baseline.[34] These studies used intensive tracking on random subsamples of 38.5% in 2012, and 36% in 2017, selected each year among all those not found during the regular tracking phase. The effective response rates were 84% in 2012 and 87% in 2017, with attrition significantly

[34] Following Blattman, Fiala and Martinez (2020) we use the standardized income family index which is composed of three measures: monthly net earnings, nondurable consumption, and durable assets.

higher in the control group. To account for this unbalanced attrition, the authors used IPW as their main specification.

Tables 10 and 11 therefore compare the ITT and WLS estimates using the standard IPW with estimates with the new IPW method.[35] To estimate the IPW, column 2 (in Table 10) and columns 2 and 3 (in Table 11) use the set of covariates used in Blattman, Fiala and Martinez (2020) to model sample selection on the full sample. Columns 3 and 4 (in Table 10) and columns 4 and 5 (in Table 11) use a vector of covariates selected following the steps in Section 5.2. Table 10 shows that new IPW estimates on income 4 years after the baseline are larger than estimates applying standard IPW after complete tracking, which themselves are larger than the ITT estimates. Comparing estimates in the top and bottom panel shows that estimates on those found during regular tracking also underestimates the effects found when using the complete tracking sample. Results from the 9 years evaluation in Table 11, on the other hand, show that ITT estimates (column 1) or IPW estimates only using those found during regular tracking (comparing top and bottom panel in columns 2 and 5) leads to an overestimate. And point estimates become even smaller and close to zero using the new IPW. Hence both in the 4-year and the 9-year study, the correction introduced by overweighting those that are hard to find using the new IPW is larger than that of the standard IPW, making a difference for final results.

Given that the reason and the selective nature of attrition, as well as the extent to which it interacts with possible heterogeneous treatment effects, will differ between studies, the new IPW approach proposed in this paper will logically not always lead to a sizable correction as compared to the ITT estimates. Overall we conclude that using the information from the intensive tracking phase to overweight observations that were hard to find can help correct for differential selective attrition, and that this may be particularly valuable when studying interventions which can directly affect the decision to migrate or when considering impacts on mobile cohorts, as is the case of the program

[35] Blattman, Fiala and Martinez (2020) do not report ITT estimates, given the unbalanced attrition, but they are included here for comparison purposes. ITT specifications in Tables 10 and 11 weight observations by their inverse probability of selection into intensive tracking.

studied in this paper, and the case in Blattman, Fiala and Martinez (2014) and Blattman, Fiala and Martinez (2020). This method can be used also when intensive tracking is only done on a random subsample, though the estimation of the weights will be facilitated by having a relatively large sub-sample of individuals tracked during the ITP to estimate the model of attrition.

8 Conclusion

Attrition can affect external and internal validity of any impact evaluation, and this is particularly relevant for studies involving mobile populations in developing countries. This paper analyzes attrition bias in a randomized experiment with balanced attrition rates and shows the sensitivity of the ITT estimates to different assumptions regarding attrition and related data collection strategies. We use a 10-year longitudinal survey with an attrition rate of 10 percent, balanced between treatment groups and collected to estimate the impact of a CCT on education and labor market outcomes. Sample selection is driven by observable characteristics and those found during an intensive tracking phase are more similar to those not found, a result that mirrors findings from other longitudinal studies. Based on this insight, we propose a new approach to correct for attrition bias.

Building on the literature that proposes attrition corrections through reweighting, we assign a weight equal to one to those found during a regular tracking phase, and we predict the probability to be found for those found during an intensive tracking phase and those not found. We compare estimates using these new inverse probability weights, with ITT estimates that do not account for attrition, and with other methods to correct for attrition bias, including regression adjustments, standard WLS estimates and bounds. We show that following regular tracking practices similar to those in most empirical studies would have led to substantial overestimates, and that inferences from regular attrition correction methods would not necessarily have helped to predict the direction of the bias, possibly due to violation of the monotonicity assumption and heterogeneous treatment effects. Similarly, collecting proxy information did not entirely solve the attrition bias problem.

Conducting intensive tracking reduces attrition bias leading to more robust estimates and allows accounting for heterogeneous treatment effects. In our case study, it also increases the share of non-compliers in the final sample, which partially explains the difference in treatment effects at different stages of the tracking process. We find that not including those found during the ITP leads to an overestimate of the ITT effects on years of schooling by about 36 and on off-farm employment by 45 percent. The results highlight the importance of studying attrition bias even in projects with low and balanced attrition rates.

As opportunities for long-term follow-ups of RCTs increase, the trade-off between tracking costs and attrition bias is likely to become relevant for an increasing number of studies. We illustrate the costs of in-person tracking which are indeed non-negligible. Yet the evidence in this paper also suggests it can be hard to predict the direction of attrition bias without such tracking. The paper further shows that having data on a subset of individuals that was hard to find can help correct for attrition bias. In this light, note that the method proposed in this paper can also be applied if only a random subset is tracked intensively, as is done in some recent studies, as long as the sample size of those sampled to be intensively tracked is large enough.

The costs and benefits involved in intensively tracking a random subset or the full sample of those missing after a regular tracking phase will, more generally, be a function of a number of factors. Intensively tracking more individuals will become more important when there are many possible factors driving selective attrition in a given context (and hence the set of covariates to include in the selection estimation is large), when there are reasons to believe treatment heterogeneity is large, or when power calculations indicate that minimizing attrition is needed to have sufficient precision of the estimates. On the other hand, when costs of tracking each individual missing after the regular tracking phase are high, random tracking a subsample can become more attractive than tracking all. Tracking costs will depend on a number of context-specific factors, including the spatial dispersion of migrants and the related costs of locating an individual, conditional on having located others; and on the technology and accessibility that can help facilitate tracking. Phone surveys can limit some of these costs, and depending on the primary outcomes of interest, can offer an alternative to obtain high tracking rates for reasonable costs. These different considerations can help provide guidance for intensive tracking decisions. Using the information from hard-to-find individuals to calculate attrition-weights can then allow to get more mileage out of such intensive tracking, and as such can help improve the trade-off between costs and potential bias.

References

Ahrens, Achim, Christian B Hansen, and Mark E Schaffer. 2020. "lassopack: Model Selection and Prediction with Regularized Regression in Stata." *The Stata Journal*, 20(1): 176–235.

Alderman, Harold, Jere Behrman, Hans-Peter Kohler, John A Maluccio, and Susan Watkins. 2001. "Attrition in Longitudinal Household Survey Data: Some Tests from Three Developing Countries." *Demographic Research*, 5: 79–124.

Athey, Susan, and Guido Imbens. 2017. "The Econometrics of Randomized Experiments." *Banerjee, A. and E. Duflo (eds.) Handbook of Economic Field Experiments*, Volume 1: 73–140. Elsevier.

Baird, Sarah, Joan Hamory, and Edward Miguel. 2008. "Tracking, Attrition, and Data Quality in the Kenyan Life Panel Survey Round 1 (KLPS-1)." *University of California CIDER Working Paper*, C08-151.

Baird, Sarah, Joan Hamory Hicks, Michael Kremer, and Edward Miguel. 2016. "Worms at Work: Long-run Impacts of a Child Health Investment." *The Quarterly Journal of Economics*, 131(4): 1637–1680.

Barham, Tania, Karen Macours, and John A Maluccio. 2017. "Are Conditional Cash Transfers Fulfilling Their Promise? Schooling, Learning, and Earnings After 10 Years." *CEPR Discussion Paper*, 11937.

Beegle, Kathleen, Joachim De Weerdt, and Stefan Dercon. 2011. "Migration and Economic Mobility in Tanzania: Evidence from a Tracking Survey." *Review of Economics and Statistics*, 93(3): 1010–1033.

Behaghel, Luc, Bruno Crépon, Marc Gurgand, and Thomas Le Barbanchon. 2015. "Please Call Again: Correcting Non-Response Bias in Treatment Effect Models." *Review of Economics and Statistics*, 97(5): 1070–1080.

Behrman, Jere R, Susan W Parker, and Peter Todd. 2009. "Medium Term Impacts of the Oportunidades Conditional Cash Transfer Program on Rural Youth in Mexico." *Poverty, Inequality and Policy in Latin America, Cambridge, MA: MIT Press*, 219–70.

Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen. 2014. "Inference on Treatment Effects after Selection Among High-dimensional Controls." *The Review of Economic Studies*, 81(2): 608–650.

Blattman, Christopher, Mathilde Emeriau, and Nathan Fiala. 2018. "Do Anti-Poverty Programs Sway Voters? Experimental Evidence from Uganda." *The Review of Economics and Statistics*, 100(5): 891–905.

Blattman, Christopher, Nathan Fiala, and Sebastian Martinez. 2014. "Generating Skilled Self-Employment in Developing Countries: Experimental Evidence from Uganda." *The Quarterly Journal of Economics*, 129(2): 697–752.

Blattman, Christopher, Nathan Fiala, and Sebastian Martinez. 2020. "The long term impacts of grants on poverty: 9-year evidence from Uganda's Youth Opportunities Program." *American Economic Review: Insights*, 2(3): 287–304.

Deaton, Angus, and Nancy Cartwright. 2016. "Understanding and Misunderstanding Randomized Controlled Trials." *NBER Working paper*, No.w22595.

DiNardo, John, Justin McCrary, and Lisa Sanbonmatsu. 2006. "Constructive Proposals for Dealing with Attrition: An Empiricial Example." *NBER Working Paper*.

Dinkelman, Taryn, and Claudia Martínez A. 2014. "Investing in Schooling in Chile: The Role of Information About Financial Aid for Higher Education." *Review of Economics and Statistics*, 96(2): 244–257.

Doyle, Orla, Colm Harmon, James J Heckman, Caitriona Logue, and Seong Hyeok Moon. 2016. "Early Skill Formation and the Efficiency of Parental Investment: a Randomized Controlled Trial of Home Visiting." *Labour Economics*. http://dx.doi.org/10.1016/j.labeco.2016.11.002.

Duflo, Esther, Pascaline Dupas, and Michael Kremer. 2015. "Education, HIV, and Early Fertility: Experimental Evidence from Kenya." *American Economic Review*, 105(9): 2757–97.

Duflo, Esther, Pascaline Dupas, and Michael Kremer. 2017. "The Impact of Free Secondary Education: Experimental Evidence from Ghana." *Massachusetts Institute of Technology Working Paper Cambridge, MA.*

Duflo, Esther, Rema Hanna, and Stephen P Ryan. 2012. "Incentives Work: Getting Teachers to Come to School." *The American Economic Review*, 102(4): 1241–1278.

Falaris, Evangelos M. 2003. "The Effect of Survey Attrition in Longitudinal Surveys: Evidence from Peru, Cote d'Ivoire and Vietnam." *Journal of Development Economics*, 70: 133–157.

Filmer, Deon, and Lant H Pritchett. 2001. "Estimating Wealth Effects Without Expenditure Data -or Tears: an Application to Educational Enrollments in States of India." *Demography*, 38(1): 115–132.

Fitzgerald, John, Peter Gottschalk, and Robert A Moffitt. 1998. "An Analyisis of Sample Attrition in Panel Data: The Michigan Panel Study of Income Dynamics." *Journal of Human Resources*, 33(2): 251–299.

Fitzsimons, Emla, Bansi Malde, Alice Mesnard, and Marcos Vera-Hernandez. 2016. "Nutrition, Information and Household Behavior: Experimental Evidence from Malawi." *Journal of Development Economics*, 122: 113 – 126.

Flores, Rafael, and John A Maluccio. 2005. "Impact Evaluation of a Conditional Cash Transfer Program: The Nicaraguan Red de Protección Social." *International Food Policy Research Institute, Discussion Paper N184.*

Fuwa, Nobuhiko. 2011. "Should We Track Migrant Households When Collecting Household Panel Data? Household Relocation, Economic Mobility, and Attrition Biases in the Rural Philippines." *American Journal of Agricultural Economics*, 93(1): 56–82.

Heckman, James J. 1979. "Sample Selection Bias as a Specification Error." *Econometrica*, 47(1): 53–161.

Hill, Daniel H, and Robert J Willis. 2001. "Reducing Panel Attrition: A Search for Effective Policy Instruments." *Journal of Human Resources*, 416–438.

Hirshleifer, Sarojini, Karen Ortiz Becerra, and Dalia Ghanem. 2019. "Testing Attrition Bias in Field Experiments." *Working Paper.*

Horowitz, Joel L, and Charles F Manski. 2000. "Nonparametric Analysis of Randomized Experiments With Missing Covariate and Outcome Data." *Journal of the American Statistical Association*, 95(449): 77–84.

Hull, Peter. 2015. "IsoLATEing: Identifying Counterfactual-Specific Treatment Effects with Cross-Stratum Comparisons." *Working Paper.*

Imbens, Guido W. 2015. "Matching methods in practice: Three examples." *Journal of Human Resources*, 50(2): 373–419.

Jensen, Robert. 2010. "The (Perceived) Returns to Education and the Demand for Schooling." *The Quarterly Journal of Economics*, 125(2): 515–548.

Kling, Jeffrey R, Jeffrey B Liebman, and Lawrence F Katz. 2007. "Experimental Analysis of Neighborhood Effects." *Econometrica*, 75(1): 83–119.

Kremer, Michael, Edward Miguel, and Rebecca Thornton. 2009. "Incentives to Learn." *The Review of Economics and Statistics*, 91(3): 437–456.

Lee, David S. 2002. "Trimming for Bounds on Treatment Effects With Missing Outcomes." *NBER Working Paper*, 0277.

Lee, David S. 2009. "Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects." *Review of Economic Studies*, 76(3): 1071–1102.

Macours, Karen, Norbert Schady, and Renos Vakis. 2012. "Cash Transfers, Behavioral Changes, and Cognitive Development in Early Childhood: Evidence from a Randomized Experiment." *American Economic Journal: Applied Economics*, 4(2): 247–273.

Maluccio, John A. 2004. "Using Quality of Interview Information to Assess Nonrandom Attrition Bias in Developing-Country Panel Data." *Review of Development Economics*, 8(1): 91–109.

Molina-Millán, Teresa, and Karen Macours. 2017. "Attrition in Randomized Control Trials: Using Tracking Information to Correct Bias." *CEPR Discussion Paper 11962*.

Orr, Larry, Judith Feins, Robin Jacob, Eric Beecroft, Lisa Sanbonmatsu, Lawrence F Katz, Jeffrey B Liebman, and Jeffrey R Kling. 2003. "Moving to Opportunity Interim Impacts Evaluation." Washington, DC: U.S. Department of Housing and Urban Development.

Robins, James M, Andrea Rotnitzky, and Lue Ping Zhao. 1995. "Analysis of Semiparametric Regression Models for Repeated Outcomes in the Presence of Missing Data." *Journal of the American Statistical Association*, 90(429): 106–121.

Rosenzweig, Mark R. 2003. "Payoffs from Panels in Low-Income Countries: Economic Development and Economic Mobility." *American Economic Review*, 93(2): 112–117.

Rubin, Donald B. 1987. *Multiple Imputation for Nonresponse in Surveys.* Wiley.

Thomas, Duncan, Elizabeth Frankenberg, and James P Smith. 2001. "Lost but Not Forgotten: Attrition and Follow-Up in the Indonesia Family Life Survey." *Journal of Human Resources*, 36: 556–592.

Thomas, Duncan, Firman Witoelar, Elizabeth Frankenberg, Bondan Sikoki, John Strauss, Cecep Sumantri, and Wayan Suriastini. 2012. "Cutting the Costs of Attrition: Results from the Indonesia Family Life Survey." *Journal of Development Economics*, 98(1): 108–123.

Tibshirani, Robert. 1996. "Regression Shrinkage and Selection via the LASSO." *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1): 267–288.

Velasquez, Andrea, Maria E Genoni, Luis Rubalcava, Graciela Teruel, and Duncan Thomas. 2010. "Attrition in Longitudinal Surveys: Evidence from the Mexican Family Life Survey." *Northeast Universities Development Consortium Conference (Proceedings).*

Wooldridge, Jeffrey M. 2002*a*. *Econometric Analysis of Cross Section and Panel Data.* The MIT press.

Wooldridge, Jeffrey M. 2002*b*. "Inverse Probability Weighted M-Estimators for Sample Selection, Attrition, and Stratification." *Portuguese Economic Journal*, 1(2): 117–139.

Zabel, Jeffrey E. 1998. "An Analysis of Attrition in the Panel Study of Income Dynamics and the Survey of Income and Program Participation With an Application to a Model of Labor Market Behavior." *Journal of Human Resources*, 33(2): 479–506.

Tables and Figures

Table 1: Final tracking status in 2010

	Early treatment	Late treatment	Total
Found	89.12	90.55	89.81
Found and interviewed	87.76	89.27	88.49
Found, dead	1.36	1.27	1.32
Not found	10.88	9.45	10.19
Not found, but info from other member	5.27	6.36	5.80
Complete missing	5.61	2.91	4.31
Refuse	0.00	0.18	0.09
Observations	588	550	1,138

Table 2: Response rates by tracking phase and treatment group

	Mean		
	Early treatment	Late treatment	Diff. means (s.e.)
By tracking phase			
CTP Sample	0.882	0.897	-0.012
			(0.025)
RTP Sample	0.706	0.755	-0.044
			(0.038)
ITP Sample	0.597	0.581	0.054
			(0.073)
By observable subgroup (after RTP)			
Mother without education	0.814	0.678	-0.136***
			(0.035)
Household size above median	0.774	0.658	-0.116**
			(0.047)
Address in hacienda	0.651	0.396	-0.255***
			(0.066)
Log expenditures pc below median	0.769	0.640	-0.129***
			(0.045)

Note: See Appendix A for subgroups variable definitions. Attrition rates by tracking phase remain balanced if we include those who were found. Estimates for differences control for stratification fixed effects. Standard errors are clustered at locality level. Robust s.e. in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 3: Intent-To-Treat

	Complete tracking phase		Regular tracking phase		Intensive tracking phase	
	Strata, age & education f.e.	+unbalanced & baseline controls	Strata, age & education f.e.	+unbalanced & baseline controls	Strata, age & education f.e.	+unbalanced & baseline controls
	(1)	(2)	(3)	(4)	(5)	(6)
Grades attained						
Early-Treatment	0.427**	0.319*	0.613***	0.350*	-0.481	0.053
	(0.17)	(0.16)	(0.20)	(0.18)	(0.37)	(0.50)
Outcome mean	5.45	5.45	5.39	5.39	5.81	5.81
R-squared	0.42	0.46	0.43	0.50	0.48	0.55
Comparing coefficients at different stages during the tracking process: P-values						
$ET^{CTP} - ET^{RTP} = 0$			0.0361	0.7219		
$ET^{CTP} - ET^{ITP} = 0$					0.0149	0.5342
$ET^{RTP} - ET^{ITP} = 0$					0.0143	0.5417
Off-Farm Employment						
Early-Treatment	0.061*	0.055**	0.088**	0.083***	-0.097*	-0.125**
	(0.03)	(0.02)	(0.03)	(0.03)	(0.06)	(0.06)
Outcome mean	0.83	0.83	0.81	0.81	0.92	0.92
R-squared	0.04	0.08	0.05	0.11	0.24	0.37
Comparing coefficients at different stages during the tracking process: P-values						
$ET^{CTP} - ET^{RTP} = 0$			0.0076	0.0054		
$ET^{CTP} - ET^{ITP} = 0$					0.0030	0.001
$ET^{RTP} - ET^{ITP} = 0$					0.0027	0.001
Obs.	1006	1006	827	827	179	179

Note: Estimates based on OLS regressions using Equation 1. The 1st model includes strata fixed effects, 3 monthly age fixed effects and set of dummies indicating whether individual had 1, 2, 3 or at least 4 years of education at baseline; additionally the 2nd model includes baseline controls for distance to school, number of children 0-8 and 9-12 in the household, estimated per capita consumption and estimated per capita consumption squared, regional fixed effects, and a vector of covariates that was off-balance after the tracking phase considered in the estimation. After complete tracking these off-balance controls are whether the individual was working, the number of individuals with family ties in village, the village population size and a productive asset index (2nd principal component of household wealth). After regular tracking the off-balance controls are mother with no education, mother with at least three years of education, the individual is son of the household head, the number of children of the household head and female household head. And in the regression on the sample targeted during ITP the off-balance controls are the number of individuals with family ties in village, the village population size and a productive asset index (2nd principal component of household wealth). Standard errors are clustered at locality level. Robust standard errors are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 4: Baseline characteristics (2000) by tracking phase

	Complete track	Regular track	Intensive track	Attritors	Diff.	Diff.
	(1)	(2)	(3)	(4)	(1)-(4)	(3)-(4)
Proxy's of permanent residence in village						
Own house (= 1)	0.844	0.861	0.764	0.709	0.135*	0.055
					(0.068)	(0.092)
House is obtained in exchange for service/labor (= 1)	0.069	0.047	0.165	0.167	-0.098	-0.001
					(0.059)	(0.082)
Address in hacienda (= 1)	0.134	0.115	0.219	0.217	-0.083	0.003
					(0.064)	(0.087)
Address in hacienda & house rented (= 1)	0.063	0.048	0.130	0.152	-0.089	-0.022
					(0.054)	(0.088)
Social Capital						
Family network size (individuals)	85.328	88.544	70.901	44.078	41.250***	26.823***
					(6.807)	(8.008)
Population size village	485.121	481.333	502.115	514.407	-29.286	-12.292
					(48.301)	(58.114)
Village Characteristics						
Village affected by hurrican Mitch (= 1)	0.871	0.877	0.847	0.902	-0.031	-0.055
					(0.032)	(0.044)
Altitude of village	628.461	630.422	619.662	635.981	-7.521	-16.320
					(28.461)	(33.410)
Village in coffee producing region (= 1)	0.783	0.784	0.777	0.787	-0.005	-0.010
					(0.053)	(0.068)
Distance to night light (meters)	17096.884	16568.363	19467.578	21805.355	-4708.470***	-2337.777**
					(1141.060)	(1009.574)
Live in Tuma region (= 1)	0.410	0.388	0.506	0.728	-0.318***	-0.222***
					(0.047)	(0.054)
Live in Madriz region (= 1)	0.203	0.213	0.161	0.064	0.140**	0.097**
					(0.052)	(0.042)
Household Characteristics: Economic activities & ***Assets***						
Household head main occupation is ag. (= 1)	0.823	0.834	0.774	0.733	0.089	0.040
					(0.057)	(0.084)
Size of landholdings ('000 sq meters)	16.260	16.749	14.066	18.584	-2.325	-4.518
					(5.740)	(6.629)
Number of parcels of land	0.964	0.979	0.899	0.711	0.253***	0.189
					(0.075)	(0.123)
Log predicted expenditures (pc)	7.727	7.737	7.679	7.768	-0.041	-0.089
					(0.052)	(0.063)
Wealth index - housing characteristics	0.048	0.061	-0.007	0.369	-0.321	-0.377
					(0.268)	(0.292)
Wealth index - productive assets	-0.005	0.028	-0.150	-0.421	0.416***	0.271*
					(0.121)	(0.150)
Wealth index - other assets	-0.038	-0.060	0.061	0.263	-0.301**	-0.202
					(0.116)	(0.126)

Table 4: Baseline characteristics (2000) by tracking phase (Continue)

	Complete Track.	Regular Track.	Intensive Track.	Attritors	Diff.	Diff.
	(1)	(2)	(3)	(4)	(1)-(4)	(3)-(4)
Household Characteristics: Demographics						
Father not living in same household (= 1)	0.188	0.175	0.246	0.369	-0.182***	-0.123*
					(0.051)	(0.063)
Mother not living in same household (= 1)	0.076	0.068	0.109	0.141	-0.065	-0.032
					(0.053)	(0.059)
Child of household head (= 1)	0.870	0.871	0.865	0.811	0.059	0.054
					(0.056)	(0.061)
Number of children of household head	5.027	5.022	5.048	4.361	0.666**	0.687*
					(0.314)	(0.397)
Female household head (= 1)	0.099	0.097	0.112	0.186	-0.086**	-0.074*
					(0.035)	(0.040)
Age of household head	44.824	44.519	46.191	42.014	2.810*	4.178**
					(1.665)	(1.909)
Number of household members	8.260	8.206	8.502	8.185	0.076	0.317
					(0.475)	(0.537)
Nuclear household (= 1)	0.618	0.636	0.533	0.540	0.077	-0.007
					(0.065)	(0.077)
Multigenerational household (= 1)	0.264	0.249	0.331	0.203	0.061	0.128*
					(0.056)	(0.069)
Other household structure (= 1)	0.118	0.114	0.136	0.257	-0.138**	-0.121*
					(0.062)	(0.062)
Number of children aged 0-8	2.073	2.038	2.229	2.451	-0.378	-0.222
					(0.276)	(0.418)
Number of children aged 9-12	1.773	1.765	1.810	1.861	-0.088	-0.051
					(0.115)	(0.135)
Household Characteristics: Education						
Distance to nearest school (minutes)	24.561	24.419	25.194	27.663	-3.103	-2.469
					(2.155)	(3.216)
Mother no grades attained (= 1)	0.477	0.459	0.557	0.271	0.206***	0.286***
					(0.074)	(0.103)
Mother 3 plus grades attained (= 1)	0.333	0.343	0.292	0.376	-0.043	-0.084
					(0.091)	(0.104)
Household head no grades attained (= 1)	0.517	0.509	0.553	0.553	-0.036	-0.000
					(0.072)	(0.092)
Household head 3 plus grades attained (= 1)	0.289	0.296	0.254	0.258	0.031	-0.004
					(0.053)	(0.054)
Individual Characteristics						
Age at start of transfer in months	10.977	10.973	10.997	11.418	-0.441***	-0.421***
					(0.121)	(0.136)
Highest grade attained	1.190	1.195	1.167	1.171	0.019	-0.004
					(0.275)	(0.305)
No grades attained (= 1)	0.443	0.441	0.452	0.554	-0.111	-0.102
					(0.082)	(0.107)
Worked in last week (= 1)	0.186	0.181	0.209	0.217	-0.031	-0.009
					(0.043)	(0.052)
Participated in some economic activity (= 1)	0.257	0.252	0.276	0.291	-0.034	-0.014
					(0.050)	(0.053)
Obs.	1,022	842	183	116		

Note: See Appendix A for definitions of all the variables. Standard errors are clustered at locality level. Robust s.e. in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 5: Compliance, baseline attrition and migration (2000) by tracking phase

	Complete Track.	Regular Track.	Intensive Track.	Attritors	Diff.	Diff.
	(1)	(2)	(3)	(4)	(1)-(4)	(3)-(4)
Received any transfer (= 1)	0.941	0.952	0.893	0.836	0.105***	0.057*
					(0.029)	(0.030)
Variables Indicating Very Early Attrition						
Probability of attrition prior to program start in locality	0.086	0.085	0.093	0.109	-0.022	-0.015
					(0.016)	(0.016)
Nobody of target sample attrited before program start	0.478	0.497	0.391	0.259	0.219***	0.131**
					(0.074)	(0.061)
Migration						
Permanent Migration	0.322	0.190	0.913	1.000	-0.678***	-0.087***
					(0.026)	(0.026)
Migration outside original municipality	0.098	0.006	0.513	0.709	-0.610***	-0.196**
					(0.076)	(0.090)
Migration outside original department	0.066	0.002	0.353	0.509	-0.443***	-0.156*
					(0.073)	(0.085)
Obs.	1,022	842	183	116		

Note: See appendix A for definitions of all the variables. Standard errors are clustered at locality level. Robust s.e. in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 6: Weighted Least Squares-Correcting for sample selection

	New IPW		Standard IPW	
	Strata, age & education f.e.	+baseline controls	Strata, age & education f.e.	+baseline controls
	(1)	(2)	(3)	(4)
Grades attained				
Complete tracking phase sample				
Early-Treatment	0.363*	0.317*	0.392**	0.282*
	(0.18)	(0.16)	(0.19)	(0.16)
Outcome mean	5.45	5.45	5.45	5.45
R-squared	0.42	0.46	0.41	0.46
Obs.	1006	1006	1006	1006
Regular tracking phase sample				
Early-Treatment			0.655***	0.362**
			(0.21)	(0.17)
Outcome mean			5.39	5.39
R-squared			0.43	0.49
Obs.			826	826
Off-farm employment				
Complete tracking phase sample				
Early-Treatment	0.061*	0.054**	0.063**	0.055**
	(0.03)	(0.02)	(0.03)	(0.02)
Outcome mean	0.83	0.83	0.83	0.83
R-squared	0.05	0.09	0.04	0.09
Obs.	1006	1006	1006	1006
Regular tracking phase sample				
Early-Treatment			0.086**	0.076***
			(0.04)	(0.03)
Outcome mean			0.81	0.81
R-squared			0.05	0.11
Obs.			827	827

Note: Estimates based on WLS regressions using new IPW (columns 1 and 2) and standard IPW (columns 3 and 4). See Section 2 for details. The 1st model includes strata fixed effects, 3 monthly age fixed effects and set of dummies indicating whether individual had 1, 2, 3 or at least 4 years of education at baseline; additionally the 2nd model includes baseline controls for distance to school, number of children 0-8 and 9-12 in the household, estimated per capita consumption and estimated per capita consumption squared, regional fixed effects and a vector of covariates that was off-balance after the tracking phase considered in the estimation. After complete tracking these off-balance baseline controls are whether the individual was working, the number of individuals with family ties in village, the village population size and a productive asset index (2nd principal component of household wealth). After regular tracking the off-balance controls are mother with no education, mother with at least three years of education, the individual is son of the household head, the number of children of the household head and female household head. And in the regression on the sample targeted during ITP the off-balance controls are the number of individuals with family ties in village, the village population size and a productive asset index (2nd principal component of household wealth). Standard errors are clustered at locality level. Robust standard errors are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 7: Lower and upper bounds

	Complete tracking phase			Regular tracking phase		
	Benchmark	Lower	Upper	Benchmark	Lower	Upper
	(1)	(2)	(3)	(4)	(5)	(6)
			Grades attained			
Cases 1 and 2: Manski bounds and bounds *à la* Kling						
OLS	0.427**			0.613***		
	(0.17)			(0.20)		
Worst-Best Scenario		-1.181***	2.132***		-3.349***	4.685***
		(0.27)	(0.26)		(0.30)	(0.34)
Mean ± 0.75 s.d		0.001	1.029***		-0.425**	2.066***
		(0.16)	(0.16)		(0.18)	(0.20)
Mean ± 0.50 s.d		0.172	0.858***		-0.010	1.651***
		(0.16)	(0.15)		(0.17)	(0.19)
Mean ± 0.25 s.d		0.344**	0.686***		0.405**	1.236***
		(0.15)	(0.15)		(0.17)	(0.18)
Case 3: Lee Bounds						
Lee Bounds		0.463**	0.698***		0.471	1.32***
		[0.234]	[0.246]		[0.36]	[0.36]
Obs.		1122	1122		1122	1122
%-trimmed		0.017	0.017		0.068	0.068
			Off-farm employment			
Cases 1 and 2: Manski bounds and bounds *à la* Kling						
OLS	0.061*			0.088**		
	(0.03)			(0.03)		
Worst-Best Scenario		-0.060*	0.161***		-0.219***	0.316***
		(0.03)	(0.03)		(0.04)	(0.03)
Mean ± 0.75 s.d		0.005	0.119***		-0.050*	0.227***
		(0.03)	(0.03)		(0.03)	(0.03)
Mean ± 0.50 s.d		0.024	0.100***		-0.004	0.181***
		(0.03)	(0.03)		(0.03)	(0.03)
Mean ± 0.25 s.d		0.043	0.081***		0.042	0.135***
		(0.03)	(0.03)		(0.03)	(0.03)
Case 3: Lee Bounds						
Lee Bounds		0.044	0.062***		0.027	0.101***
		[0.029]	[0.023]		[0.041]	[0.027]
Obs.		1122	1122		1122	1122
%-trimmed		0.017	0.017		0.068	0.068

Note: The sample used includes 1,006 individuals found and from whom we do have information on grades attained and off-farm employment in 2010, and 116 individuals not found. The sample does not include 15 deceased individuals found in 2010. All the regressions include strata fixed effects, 3 monthly age fixed effects and set of dummies indicating whether individual had 1,2,3 or at least 4 years of education at baseline. Manski Bounds and Bounds *à la* Kling: Standard errors are clustered at locality level. Robust s.e. in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 8: ITT: Correcting for attrition using proxy information on attritors

	CTP	RTP	RTP+ITP sample with double information		CTP sample + Attritors	RTP sample + ITP & Attritors
	Benchmark	Benchmark	Self-report	Proxy-Inform.	Proxy-Inform.	Proxy-Inform.
	(1)	(2)	(3)	(4)	(5)	(6)
			Grades attained			
Early-Treatment	0.427**	0.613***	0.523***	0.512***	0.424**	0.516***
	(0.17)	(0.19)	(0.19)	(0.19)	(0.17)	(0.18)
Outcome Mean	5.78	5.88	5.86	5.84	5.68	5.75
Obs.	1006	826	881	881	1072	934
R-squared	0.42	0.43	0.42	0.42	0.41	0.42
			Off-farm employment			
Early-Treatment	0.061**	0.088***	0.086**	0.092***	0.058*	0.090**
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Outcome Mean	0.86	0.86	0.87	0.86	0.86	0.85
Obs.	1006	827	881	881	1071	933
R-squared	0.04	0.05	0.05	0.05	0.04	0.05

Note: All the regressions include strata fixed effects, 3 monthly age fixed effects and set of dummies indicating whether individual had 1,2,3 or at least 4 years of education at baseline. Standard errors are clustered at locality level. Robust s.e. in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 9: OLS and WLS - Macours, Schady and Vakis (2012)

	Cognitive & socio-emotional index		
	ITT	New IPW	Standard IPW
	(1)	(2)	(3)
Complete tracking phase sample			
ITT	0.084**	0.084**	0.079*
	(0.04)	(0.04)	(0.04)
R-squared	0.43	0.43	0.43
Obs.	2771	2771	2769
Regular tracking phase sample			
ITT	0.114**		0.113**
	(0.05)		(0.05)
R-squared	0.49		0.49
Obs.	2029		2027

Note: All specifications include strata fixed effects, 6 monthly age fixed effects and an indicator variable equal to 1 if the child is female. Column 1 presents estimates based on OLS regression. Column 2 presents estimates based on WLS regressions with weights predicted by estimating a selection model using the intensive tracking phase sample. Column 3 reports WLS estimates with weights predicted by estimating a selection model on the complete tracking sample. For columns 2 and 3, a stepwise selection of covariates (as described in Section 5.2) was used for the selection models. Robust s.e. in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Standard errors are clustered at the community level.

Table 10: WLS - Blattman, Fiala and Martinez (2014)

	Standardized Income Index - Treatment effects after 4 years			
	ITT	New IPW		Standard IPW
covariates used to construct IPW →		Authors' covariates	New set covariates	New set covariates
	(1)	(2)	(3)	(4)
Complete tracking phase sample				
Assigned to treatment	0.224***	0.278***	0.251***	0.223***
	(0.05)	(0.06)	(0.05)	(0.05)
R-squared	0.23	0.26	0.24	0.23
Obs.	1868	1868	1868	1868
Regular tracking phase sample				
Assigned to treatment	0.172***			0.173***
	(0.05)			(0.05)
R-squared	0.23			0.23
Obs.	1632			1632

Note: Estimates are based on WLS regressions. All specifications control for district fixed effects and a vector of 38 baseline controls (See Table 2 in Blattman, Fiala and Martinez (2014)). Regressions in column 1 are weighted by inverse probability of selection into the ITP tracking sample. Columns 2 and 3 present WLS estimates with weights predicted by estimating a selection model using the intensive tracking phase sample. For column 2, the full set of variables used as controls were used in the selection model. For column 3, a stepwise selection of covariates (as described in Section 5.2) was used for the selection model. Column 4 reports WLS estimates with weights predicted by estimating a selection model on the complete tracking sample, with stepwise selection of covariates for the selection model. Robust s.e. in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Standard errors are clustered at the group level (of up to 5 people).

Table 11: WLS - Blattman, Fiala and Martinez (2020)

	Standardized Income Index - Treatment effects after 9 years				
	ITT	Standard IPW	New IPW		Standard IPW
covariates used to construct IPW →		Authors' covariates	Authors' covariates	New set covariates	New set covariates
	(1)	(2)	(3)	(4)	(5)
Complete tracking phase sample					
Assigned to treatment	0.092*	0.078	-0.010	0.037	0.092*
	(0.05)	(0.05)	(0.07)	(0.06)	(0.05)
R-squared	0.20	0.21	0.73	0.21	0.20
Obs.	1981	1981	1981	1981	1981
Regular tracking phase sample					
Assigned to treatment	0.160***	0.158***			0.161***
	(0.04)	(0.04)			(0.04)
R-squared	0.22	0.25			0.22
Obs.	1903	1903			1903

Note: Estimates are based on WLS regressions. All specifications control for district fixed effects and a vector of 38 baseline controls (See Table 2 in Blattman, Fiala and Martinez (2020)). Regressions in column 1 are weighted by inverse probability of selection into the ITP tracking sample. Columns 2 and 5 reports WLS estimates with weights predicted by estimating a selection model on the complete tracking sample. Columns 3 and 4 present WLS estimates with weights predicted by estimating a selection model using the intensive tracking phase sample. For columns 2 and 3, the full set of variables used as controls were used in the selection model. For columns 4 and 5, a stepwise selection of covariates (as described in Section 5.2) was used for the selection model. Robust s.e. in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Standard errors are clustered at the group level (of up to 5 people).

Figure 1: Cost analysis: ITT estimates on grades attained during the ITP

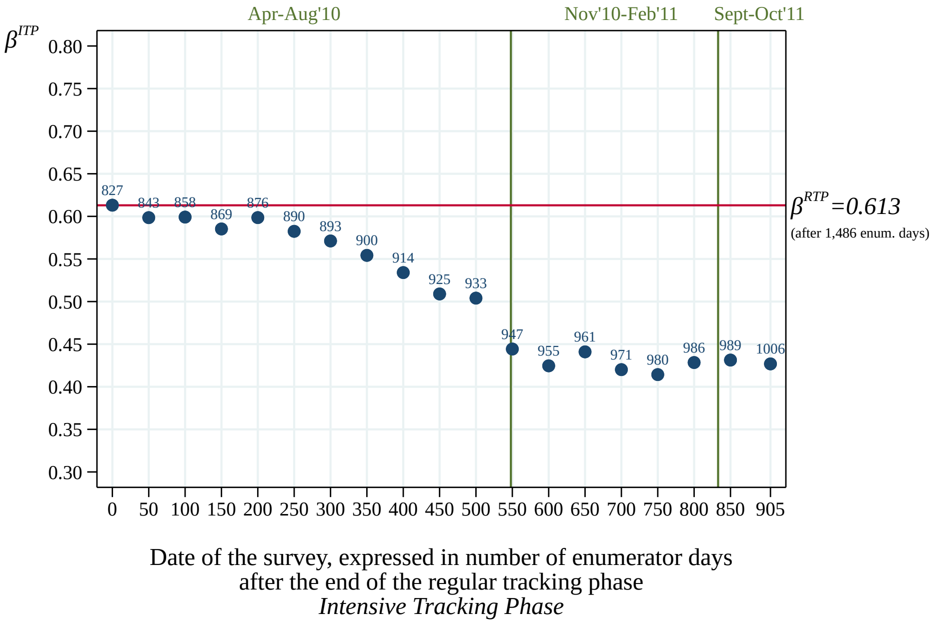
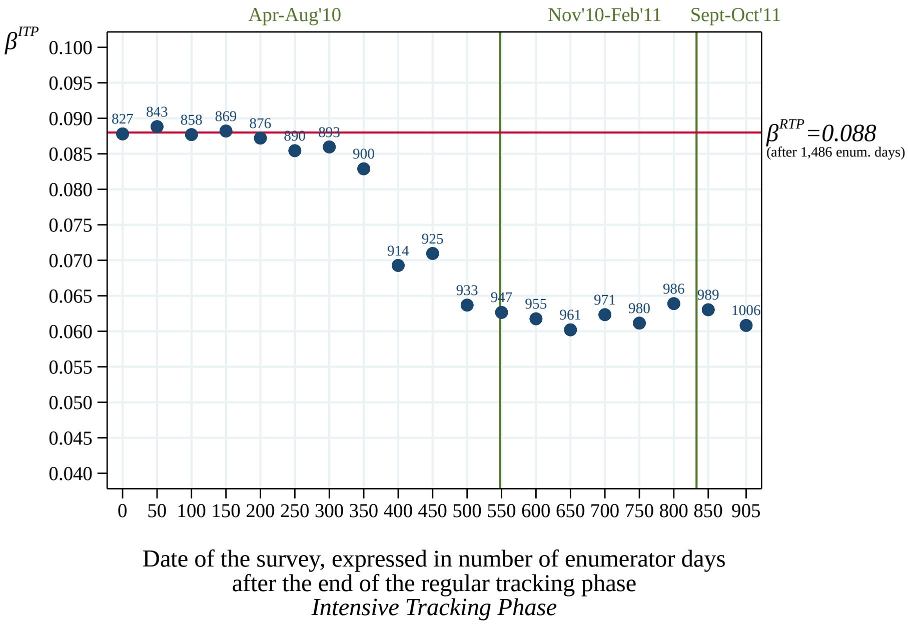


Figure 2: Cost analysis: ITT estimates on off-farm employment during the ITP



Note (Figures 1 & 2): Estimates based on OLS regressions using Equation 1. The figures plot ITT estimates (β) on the sample of respondents found during the RTP, plus those found during the ITP after x days of enumerators days. An enumerator day is defined as any working day in which the team of enumerators worked after RTP (March, 23th 2010) times the number of enumerators in the team at each date. Numbers above the dots show the number of male respondents between the ages of 9 and 12 years found until that day. All specifications include strata fixed effects, 3 monthly age fixed effects and set of dummies indicating whether individual had 1, 2, 3 or at least 4 years of education at baseline. The horizontal red line shows the value of the ITT estimate after conducting RTP. The vertical green lines mark the different phases of the ITP (see Appendix C).

Online Appendices for “Attrition in Randomized Control Trials: Using tracking information to correct bias”

January 18, 2021

A Appendix

List of variables

- **Early-Treatment:** An ITT indicator that takes value of one for children in localities randomly assigned to early treatment and zero otherwise
- **Own house** (= 1)**:** Dichotomous variable that takes value of one for households owning the house and zero otherwise
- **House is obtained in exchange for service/labor** (= 1)**:** Dichotomous variable that takes value of one for households who received the house in exchange of services and zero otherwise
- **Address in hacienda** (= 1)**:** Dichotomous variable that takes value of one for households whose address was in an large coffee plantation (*haciendas*) and zero otherwise
- **Address in hacienda** & **house rented** (= 1)**:** Dichotomous variable that takes value of one for households whose address was in an large coffee plantation (*haciendas*) and who obtained a house in exchange of services and zero otherwise
- **Family network size (individuals):** Number of individuals by village with at least one co-villager having the same last name
- **Population size village:** Number of individuals living in a village
- **Village affected by hurricane Mitch** (= 1)**:** Dichotomous variable that takes value of one for villages affected by the hurricane Mitch in 1998 and zero otherwise
- **Altitude of village:** Village's altitude measured using GPS (in meters)
- **Village in coffee producing region** (= 1)**:** Dichotomous variable that takes value of one for villages located in a coffee producing regions and zero otherwise. Source of coffee producing areas: International Center for Tropical Agriculture (CIAT)

- **Distance to night light (meters):** Linear distance in meters from household location to an area with stables night lights detected by a satellite. Source of night lights: DMSP-OLS Nighttime Lights

- **Live in Tuma region** $(= 1)$**:** Dichotomous variable that takes value of one for households located in El Tuma – La Dalia municipality and zero otherwise

- **Live in Madriz region** $(= 1)$**:** Dichotomous variable that takes value of one for households located in Madriz department and zero otherwise

- **Household head main occupation is ag.** $(= 1)$**:** Dichotomous variable that takes value of one for household heads working in agriculture and zero otherwise

- **Size of landholdings ('000 sq meters):** Household's agricultural land holdings in square meters

- **Log of size of landholdings:** Logarithm of household's agricultural land holdings

- **Number of parcels of land:** Household's number of parcels

- **Log predicted expenditures (pc):** Logarithm of predicted per capita expenditures

- **Wealth index - housing characteristics:** First Principal Component estimate capturing household characteristics (see appendix D for details)

- **Wealth index - productive assets:** Second Principal Component estimate capturing household productive assets (see appendix D for details)

- **Wealth index - other assets:** Third Principal Component estimate capturing household other assets (see appendix D for details)

- **Father not living in same household** $(= 1)$**:** Dichotomous variable that takes value of one for children not living in the same house as his father and zero otherwise

- **Mother not living in same household** $(=1)$**:** Dichotomous variable that takes value of one for children not living in the same house as his mother and zero otherwise

- **Child of household head** $(=1)$**:** Dichotomous variable that takes value of one for children whose father or mother was the household head and zero otherwise

- **Number of children of household head:** Number of household head children living in the household

- **Female household head** $(=1)$**:** Dichotomous variable that takes value of one for households with a female household head and zero otherwise

- **Age of household head:** Age of the household head

- **Number of household members:** Number of household members

- **Nuclear household** $(=1)$**:** Dichotomous variable that takes value of one for households consisting entirely of a single family nucleus and zero otherwise

- **Multigenerational household** $(=1)$**:** Dichotomous variable that takes value of one for multi-generational household and zero otherwise

- **Other household structure** $(=1)$**:** Dichotomous variable that takes value of one for extended household and zero otherwise

- **Number of children aged 0-8:** Number of household members ages 0-8

- **Number of children aged 9-12:** Number of household members ages 9-12

- **Distance to nearest school (minutes):** Distance in minutes from household location to school

- **Mother no grades attained** $(=1)$**:** Dichotomous variable that takes value of one for children whose mother had zero years of education and zero otherwise

- **Mother 3 plus grades attained** (= 1)**:** Dichotomous variable that takes value of one for children whose mother had at least three years of education and zero otherwise

- **Household head no grades attained** (= 1)**:** Dichotomous variable that takes value of one for children whose household head had zero years of education and zero otherwise

- **Household head 3 plus grades attained** (= 1)**:** Dichotomous variable that takes value of one for children whose household head had at least three years of education and zero otherwise

- **Age at start of transfer in months:** Age by November 2000

- **Highest grade attained:** Number of grades of education attained

- **No grades attained** (= 1)**:** Dichotomous variable that takes value of one for children with zero grades of education attained and zero otherwise

- **Worked in last week** (= 1)**:** Dichotomous variable that takes value of one for children who worked the week prior to the survey and zero otherwise

- **Participated in some economic activity** (= 1)**:** Dichotomous variable that takes value of one for children who participated in an economic activity the week prior to the survey and zero otherwise

- **Received any transfer:** Dichotomous variable that takes value of one for households who received any transfer from RPS and zero otherwise

- **Probability of attrition prior to program start in locality:** Attrition rate at the locality level between the Census in 2000 and the start of the program

- **Nobody of target sample attrited before program start:** Dichotomous variable that takes value of one for localities with attrition prior to start of the program equals to zero and zero otherwise

- **Permanent Migration:** Dichotomous variable that takes value of one for migrants absent for more than nine months in the last 12 and zero otherwise

- **Migration outside original municipality:** Dichotomous variable that takes value of one for migrants moving to another municipality and zero otherwise

- **Migration outside original department:** Dichotomous variable that takes value of one for migrants moving to another department and zero otherwise

Figure A1: Distribution of weights - Complete tracking sample

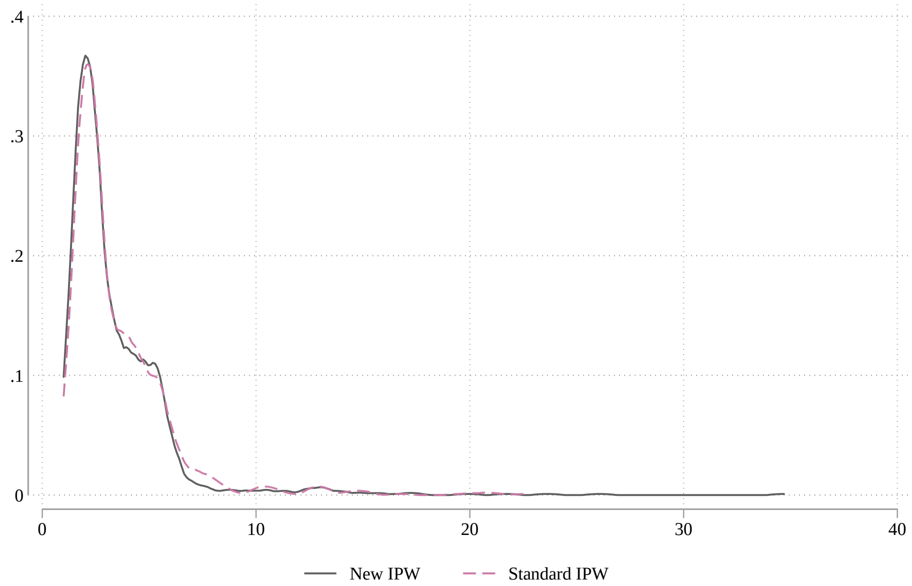


Table A1: Baseline characteristics (2000): differences between experimental groups by tracking phase

	Full baseline Sample	Complete Tracking	Regular Tracking
	ET-LT (s.e)	ET-LT (s.e)	ET-LT (s.e)
Proxy's of permanent residence in village			
Own house (= 1)	-0.052	-0.061	-0.020
	(0.043)	(0.043)	(0.046)
House is obtained in exchange for service/labor (= 1)	0.019	0.019	-0.018
	(0.039)	(0.036)	(0.032)
Address in hacienda (= 1)	-0.022	-0.0089	-0.028
	(0.058)	(0.056)	(0.058)
Address in hacienda & house rented (= 1)	0.032	0.043	0.015
	(0.033)	(0.031)	(0.031)
Social Capital			
Family network size (individuals)	23.5**	26.0**	27.8**
	(10.3)	(10.8)	(12.1)
Population size village	257.6***	256.0***	247.1***
	(88.6)	(89.3)	(95.1)
Village Characteristics			
Village affected by hurrican Mitch (= 1)	-0.062	-0.055	-0.057
	(0.050)	(0.052)	(0.051)
Altitude of village	-21.1	-22.1	-25.5
	(34.6)	(36.2)	(36.4)
Village in coffee producing region (= 1)	-0.018	-0.0084	-0.022
	(0.065)	(0.066)	(0.067)
Distance to night light (meters)	2276.2	1882.3	1276.3
	(2642.3)	(2629.1)	(2642.7)
Live in Tuma region (= 1)	0.19	0.19	0.16
	(0.15)	(0.14)	(0.14)
Live in Madriz region (= 1)	0.017	0.029	0.048
	(0.12)	(0.13)	(0.12)
Household Characteristics: Economic activities & Assets			
Household head main occupation is ag. (= 1)	-0.021	-0.020	-0.0054
	(0.038)	(0.040)	(0.046)
Size of landholdings ('000 sq meters)	-1.99	-2.41	-2.73
	(1.85)	(1.79)	(2.01)
Number of parcels of land	-0.036	-0.0075	-0.017
	(0.073)	(0.074)	(0.075)
Log predicted expenditures (pc)	-0.0089	0.0021	0.027
	(0.022)	(0.025)	(0.026)
Wealth index - housing characteristics	0.20	0.22	0.32
	(0.16)	(0.15)	(0.16)
Wealth index - productive assets	-0.26**	-0.26**	-0.21**
	(0.11)	(0.11)	(0.12)
Wealth index - other assets	-0.040	-0.061	-0.13
	(0.18)	(0.19)	(0.19)

Table A1: Baseline characteristics (2000): differences between experimental groups by tracking phase (continue)

	All Targeted Sample	Complete Tracking	Regular Tracking
	ET-LT (s.e)	ET-LT (s.e)	ET-LT (s.e)
Household Characteristics: Demographics			
Father not living in same household (= 1)	0.017	0.012	0.042
	(0.031)	(0.027)	(0.029)
Mother not living in same household (= 1)	0.011	0.013	0.029
	(0.017)	(0.017)	(0.018)
Child of household head (= 1)	-0.016	-0.017	-0.060**
	(0.024)	(0.024)	(0.024)
Number of children of household head	-0.24	-0.32	-0.51**
	(0.22)	(0.23)	(0.23)
Female household head (= 1)	0.030	0.020	0.037**
	(0.020)	(0.017)	(0.018)
Age of household head	0.34	0.41	0.52
	(0.87)	(0.81)	(0.88)
Number of household members	-0.047	-0.17	-0.21
	(0.19)	(0.22)	(0.24)
Nuclear household (= 1)	-0.017	-0.021	-0.027
	(0.040)	(0.043)	(0.044)
Multigenerational household (= 1)	-0.037	-0.015	-0.00043
	(0.035)	(0.038)	(0.035)
Other household structure (= 1)	0.054*	0.036	0.027
	(0.028)	(0.028)	(0.031)
Number of children aged 0-8	0.021	-0.098	-0.067
	(0.11)	(0.11)	(0.12)
Number of children aged 9-12	-0.036	-0.076	-0.082
	(0.058)	(0.063)	(0.071)
Household Characteristics: Education			
Distance to nearest school (minutes)	0.74	0.89	2.92
	(4.76)	(4.65)	(4.71)
Mother no grades attained (= 1)	-0.046	-0.052	-0.097**
	(0.037)	(0.042)	(0.041)
Mother 3 plus grades attained (= 1)	0.072	0.065	0.11***
	(0.044)	(0.042)	(0.038)
Household head no grades attained (= 1)	0.0049	0.018	0.0017
	(0.031)	(0.034)	(0.039)
Household head 3 plus grades attained (= 1)	0.025	0.033	0.047
	(0.029)	(0.027)	(0.031)
Individual Characteristics			
Age at start of transfer in months	-0.046*	-0.068**	-0.090*
	(0.026)	(0.031)	(0.046)
Highest grade attained	0.071	0.096	0.14
	(0.16)	(0.15)	(0.16)
No grades attained (= 1)	-0.0082	-0.019	-0.029
	(0.064)	(0.063)	(0.066)
Worked in last week (= 1)	-0.054*	-0.062**	-0.063**
	(0.027)	(0.029)	(0.028)
Participated in some economic activity (= 1)	-0.0086	-0.012	-0.015
	(0.035)	(0.034)	(0.036)
Obs.	1138	1022	841

Note: Estimates for differences control for stratification fixed effects. Standard errors are clustered at locality level. Robust s.e. in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A2: Intent-To-Treat with alternative sets of control variables

	Complete tracking phase				Regular tracking phase				Intensive tracking phase			
	Stratif. Dummies	+Age&Educ. Controls	+Unbalanced Controls	+Baseline Controls	Stratif. Dummies	+Age& Educ. Controls	+Unbalanced Controls	+Baseline Controls	Stratif. Dummies	+Age & Educ. Controls	+Unbalanced Controls	+Baseline Controls
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Grades attained												
Early-Treatment	0.593**	0.427**	0.362**	0.319*	0.865**	0.613***	0.451**	0.350*	-0.745	-0.481	-0.400	0.053
	(0.29)	(0.17)	(0.18)	(0.16)	(0.33)	(0.20)	(0.20)	(0.18)	(0.58)	(0.37)	(0.43)	(0.50)
Outcome Mean	5.45	5.45	5.45	5.45	5.39	5.39	5.39	5.39	5.81	5.81	5.81	5.81
R-squared	0.04	0.42	0.42	0.46	0.05	0.43	0.46	0.50	0.10	0.48	0.48	0.55
Comparing coefficients at different stages during the tracking process: P-values												
$ET^{CTP} - ET^{RTP} = 0$					0.0330	0.0361	0.3202	0.7219				
$ET^{CTP} - ET^{ITP} = 0$									0.0241	0.0149	0.0598	0.5342
$ET^{RTP} - ET^{ITP} = 0$									0.0231	0.0143	0.0689	0.5417
Off-Farm Employment												
Early-Treatment	0.059*	0.061*	0.056*	0.055**	0.087**	0.088**	0.081**	0.083***	-0.071	-0.097*	-0.097*	-0.125**
	(0.03)	(0.03)	(0.03)	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)	(0.05)	(0.06)	(0.06)	(0.06)
Outcome Mean	0.83	0.83	0.83	0.83	0.81	0.81	0.81	0.81	0.92	0.92	0.92	0.92
R-squared	0.02	0.04	0.05	0.08	0.03	0.05	0.08	0.11	0.04	0.24	0.27	0.37
Comparing coefficients at different stages during the tracking process: P-values												
$ET^{CTP} - ET^{RTP} = 0$					0.0041	0.0076	0.0156	0.0054				
$ET^{CTP} - ET^{ITP} = 0$									0.0093	0.0030	0.0028	0.001
$ET^{RTP} - ET^{ITP} = 0$									0.0073	0.0027	0.0025	0.001
Obs.	1006	1006	1006	1006	827	827	827	827	179	179	179	179

Note: Estimates based on OLS regressions using Equation 1. The 1st model includes only strata fixed effects; in the 2nd model we add 3 monthly age fixed effects and set of dummies indicating whether individual had 1, 2, 3 or at least 4 years of education at baseline; additionally the 3rd model includes a vector of covariates that ended up off-balance after each of the tracking phases. After complete tracking the off-balance baseline controls are whether the individual was working, the number of individuals with family ties in village, the village population size and a productive asset index (2nd principal component of household wealth). After regular tracking the off-balance controls are mother with no education, mother with at least three years of education, the individual is son of the household head, the number of children of the household head and female household head. The regression on the sample targeted during ITP the off-balance baseline are the number of individuals with family ties in village, the village population size and a productive asset index (2nd principal component of household wealth); the 4th and last specification also controls for distance to school, number of children 0-8 and 9-12 in the household, estimated per capita consumption and estimated per capita consumption squared, as well as regional fixed effects. Standard errors are clustered at locality level. Robust s.e. in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A3: Weighted Least Squares-Correcting for sample selection with alternative sets of control variables

	New IPW				Standard IPW			
	Stratif. dummies	+Age&Educ. controls	+Unbalanced controls	+Baseline controls	Stratif. dummies	+Age& Educ. controls	+Unbalanced controls	+Baseline controls
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Grades attained								
Complete Tracking Phase Sample								
Early-Treatment	0.361	0.363*	0.344*	0.317*	0.503	0.392**	0.329*	0.282*
	(0.32)	(0.18)	(0.18)	(0.16)	(0.30)	(0.19)	(0.19)	(0.16)
R-squared	0.03	0.42	0.43	0.46	0.04	0.41	0.42	0.46
Obs.	1006	1006	1006	1006	1006	1006	1006	1006
Regular Tracking Phase Sample								
Early-Treatment					0.770**	0.655***	0.531**	0.362**
					(0.32)	(0.21)	(0.21)	(0.17)
R-squared					0.04	0.43	0.46	0.49
Obs.					826	826	826	826
Off-Farm Employment								
Complete Tracking Phase Sample								
Early-Treatment	0.059**	0.061*	0.055*	0.054**	0.060**	0.063**	0.058*	0.055**
	(0.03)	(0.03)	(0.03)	(0.02)	(0.03)	(0.03)	(0.03)	(0.02)
R-squared	0.02	0.05	0.06	0.09	0.02	0.04	0.06	0.09
Obs.	1006	1006	1006	1006	1006	1006	1006	1006
Regular Tracking Phase Sample								
Early-Treatment					0.088**	0.086**	0.075**	0.076***
					(0.03)	(0.04)	(0.03)	(0.03)
R-squared					0.03	0.05	0.08	0.11
Obs.					827	827	827	827

Note: Estimates based on WLS regressions using new IPW (columns 1 to 4) and standard IPW (columns 5 to 8). See Section 2 for details. The 1st model includes only strata fixed effects; in the 2nd model we add 3 monthly age fixed effects and set of dummies indicating whether individual had 1,2,3 or at least 4 years of education at baseline; additionally the 3rd model includes a vector of covariates that ended up off-balance after each of the tracking phases. After complete tracking the off-balance baseline controls are whether the individual was working, the number of individuals with family ties in village, the village population size and a productive asset index (2nd principal component of household wealth). After regular tracking the off-balance controls are mother with no education, mother with at least three years of education, the individual is son of the household head, the number of children of the household head and female household head. The regression on the sample targeted during ITP the off-balance baseline are the number of individuals with family ties in village, the village population size and a productive asset index (2nd principal component of household wealth); the 4th and last specification also controls for distance to school, number of children 0-8 and 9-12 in the household, estimated per capita consumption and estimated per capita consumption squared, as well as regional fixed effects. Standard errors are clustered at locality level. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A4: Lower and upper bounds. Cases 1 and 2: Manski bounds and bounds *à la* Kling - Grades attained

	Complete tracking phase			Regular tracking phase		
	Benchmark	Lower	Upper	Benchmark	Lower	Upper
	(1)	(2)	(3)	(4)	(5)	(6)
			Stratification F.E.			
OLS	0.593**			0.865**		
	(0.29)			(0.33)		
Worst-Best Scenario		-1.056***	2.218***		-3.223***	4.753***
		(0.31)	(0.31)		(0.35)	(0.35)
Mean ± 0.75 s.d		0.114	1.130***		-0.319	2.154***
		(0.26)	(0.27)		(0.26)	(0.25)
Mean ± 0.50 s.d		0.283	0.961***		0.093	1.742***
		(0.26)	(0.26)		(0.25)	(0.25)
Mean ± 0.25 s.d		0.453*	0.791***		0.505**	1.330***
		(0.26)	(0.26)		(0.24)	(0.24)
			Stratif.,Age and Educ. F.E. +Unbalanced Controls			
OLS	0.426**			0.558***		
	(0.16)			(0.20)		
Worst-Best Scenario		-1.180***	2.166***		-3.355***	4.686***
		(0.26)	(0.26)		(0.31)	(0.34)
Mean ± 0.75 s.d		0.019	1.059***		-0.418**	2.076***
		(0.16)	(0.16)		(0.17)	(0.19)
Mean ± 0.50 s.d		0.193	0.885***		-0.003	1.660***
		(0.15)	(0.15)		(0.17)	(0.18)
Mean ± 0.25 s.d		0.366**	0.712***		0.413**	1.244***
		(0.15)	(0.15)		(0.16)	(0.17)
			Stratif.,Age and Educ. F.E. +Unbalanced and Baseline Controls			
OLS	0.350**			0.424**		
	(0.17)			(0.20)		
Worst-Best Scenario		-1.214***	2.140***		-3.410***	4.591***
		(0.28)	(0.28)		(0.35)	(0.35)
Mean ± 0.75 s.d		-0.009	1.033***		-0.493**	1.989***
		(0.17)	(0.17)		(0.20)	(0.19)
Mean ± 0.50 s.d		0.164	0.859***		-0.080	1.575***
		(0.16)	(0.16)		(0.18)	(0.17)
Mean ± 0.25 s.d		0.338**	0.685***		0.334*	1.161***
		(0.15)	(0.15)		(0.17)	(0.17)

Note: The sample used includes 1,006 individuals found and from whom we do have information on grades attained in 2010, and 116 individuals not found. The sample does not include 15 deceased individuals found in 2010. Standard errors are clustered at locality level. Robust s.e. in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A5: Lower and upper bounds. Cases 1 and 2: Manski bounds and bounds *à la* Kling - Off-farm employment

	Complete Tracking Phase			Regular Tracking Phase		
	Benchmark	Lower	Upper	Benchmark	Lower	Upper
	(1)	(2)	(3)	(4)	(5)	(6)
Stratification F.E.						
OLS	0.059*			0.087**		
	(0.03)			(0.03)		
Worst-Best Scenario		-0.060*	0.158***		-0.219***	0.313***
		(0.03)	(0.03)		(0.04)	(0.03)
Mean ± 0.75 s.d		0.003	0.116***		-0.051*	0.223***
		(0.03)	(0.03)		(0.03)	(0.03)
Mean ± 0.50 s.d		0.022	0.097***		-0.005	0.178***
		(0.03)	(0.03)		(0.03)	(0.03)
Mean ± 0.25 s.d		0.040	0.078***		0.040	0.132***
		(0.03)	(0.03)		(0.03)	(0.02)
Stratif.,Age and Educ. F.E. +Unbalanced Controls						
OLS	0.069**			0.097***		
	(0.03)			(0.04)		
Worst-Best Scenario		-0.055	0.168***		-0.213***	0.323***
		(0.03)	(0.03)		(0.04)	(0.03)
Mean ± 0.75 s.d		0.009	0.124***		-0.046*	0.231***
		(0.03)	(0.03)		(0.03)	(0.03)
Mean ± 0.50 s.d		0.028	0.105***		0.000	0.185***
		(0.03)	(0.03)		(0.03)	(0.03)
Mean ± 0.25 s.d		0.047*	0.086***		0.046*	0.139***
		(0.03)	(0.03)		(0.03)	(0.03)
Stratif.,Age and Educ. F.E. +Unbalanced and Baseline Controls						
OLS	0.059*			0.087**		
	(0.03)			(0.03)		
Worst-Best Scenario		-0.065*	0.158***		-0.223***	0.311***
		(0.04)	(0.03)		(0.04)	(0.03)
Mean ± 0.75 s.d		-0.002	0.114***		-0.055*	0.220***
		(0.03)	(0.03)		(0.03)	(0.03)
Mean ± 0.50 s.d		0.017	0.095***		-0.009	0.174***
		(0.03)	(0.03)		(0.03)	(0.03)
Mean ± 0.25 s.d		0.037	0.075**		0.037	0.128***
		(0.03)	(0.03)		(0.03)	(0.03)

Note: The sample used includes 1,006 individuals found and from whom we do have information on off-farm employment in 2010, and 116 individuals not found. The sample does not include 15 deceased individuals found in 2010. Standard errors are clustered at locality level. Robust s.e. in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A6: Percentage of respondents from whom we have proxy information on education and labor outcomes

	Double Information			
	Complete Tracking Phase	Regular Tracking Phase	Intensive Tracking Phase	Attritors
	Grades attained			
Early-Treatment	-0.001	-0.001	-0.042	-0.169
	(0.03)	(0.03)	(0.11)	(0.15)
Obs.	1007	827	180	116
Control Mean	0.18	0.12	0.42	0.57
	Off-Farm Employment			
Early-Treatment	0.005	0.004	-0.033	-0.173
	(0.03)	(0.03)	(0.11)	(0.15)
Obs.	1007	827	180	116
Control Mean	0.18	0.12	0.42	0.57

Note: All specifications include stratification fixed effects. Standard errors are clustered at locality level. Robust s.e. in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A7: Percentage of respondents with accurate proxy information and direction of the bias for respondents without accuracy information

	Complete Tracking Phase		Regular Tracking Phase		Intensive Tracking Phase	
	Correct Proxy Information	Bias (proxy -self report)	Correct Proxy Information	Bias (proxy -self report)	Correct Proxy Information	Bias (proxy -self report)
	(1)	(2)	(3)	(4)	(5)	(6)
	Grades attained					
Early-Treatment	0.018	-0.075	0.060	0.001	-0.034	-0.128
	(0.07)	(0.07)	(0.07)	(0.10)	(0.14)	(0.15)
Obs.	175	175	98	98	77	77
Control Mean	0.56	-0.12	0.63	-0.10	0.48	-0.13
	Off-Farm Employment					
Early-Treatment	0.055	-0.018	-0.011	-0.121	0.131	0.087
	(0.08)	(0.08)	(0.11)	(0.09)	(0.12)	(0.12)
Obs.	175	175	100	100	75	75
Control Mean	0.66	-0.31	0.60	-0.36	0.75	-0.24

Note: All specifications include stratification fixed effects. Standard errors are clustered at locality level. Robust s.e. in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

B Appendix: Compliance and correlates of attrition

B.1 Correlates of Attrition

To investigate the correlates of attrition, we consider both variables potentially correlated with the decision to migrate, and variables potentially correlated with the quality of the information regarding migration destinations that can be obtained at origin. As such, differences in socioeconomic status and village conditions, existing networks, family structures and temporary residence are likely candidates to help explain differential attrition.

First, individuals with lower socio-economic status or living in more remote areas, might be more likely to migrate in search of better opportunities elsewhere. Results in Table 4 and Table B1 confirm that young adults from households with less productive assets and land, and living in more remote areas (as measured by distance to night light), were indeed more likely to attrit.

The probability to be found 10 years after the baseline survey is likely also a function of an individual's social and family relationships. Existing networks can affect the probability and destination of migration, but also the quality of the information about potential migrants obtained in the origin villages. People with larger networks may also be more likely to return, or otherwise keep stronger ties, all of which can affect tracking success. Networks of course can also affect the outcomes of interest.

This is particularly relevant in our study as some households targeted by the program were temporary workers in large coffee plantations (haciendas), rather than permanent residents of the target localities.[36] Treatment effects for these temporary workers are likely to be different than for other household, as some of them may have moved out of the area even prior to the start of the program. Others may have been induced to stay longer than intended once they became eligible, and this presumably could be more important in the early treatment villages. As the temporary workers typically did not

[36] Such plantations, often employ a large number of workers, permanently or temporary. Those living in haciendas are households with limited access to land and with few resources, whose main income comes from the agricultural wage work for the hacienda. It is common for these workers to migrate after the peak season, in search of wage opportunities elsewhere.

have a family network in the localities of origin, they may be harder to track once they have moved on.[37] Table 4 and Table B1 confirm that proxies of permanent residence in the village and networks are indeed strongly correlated with attrition. The probability of being found is also significantly lower in areas with very early migration (as measured by the attrition between the program census and baseline survey).

Finally, the probability to be found could be correlated with the demographic characteristics of the households. About 40 percent of households at baseline do not have the typical nuclear household structures. A relatively common phenomenon, in rural Nicaragua as in many other developing countries, are children living with grandparents or other family members while their parents are working elsewhere. As those are often temporary arrangements, it can be harder to find those children 10 years later, as they may have migrated before, during or after the intervention. Results in Table 4 and Table B1 confirm that attritors were also more likely to come from female headed and smaller households and were less likely to live with their biological parents. Attritors are also about half a year older.

Table B1 further shows that there are some differences in correlates of attrition between early and late treatment groups, in particular for the demographic characteristics. That said, after the full tracking only few variables show significantly differences between the two experimental groups, and in contrast with the results after the regular tracking (see further). This is in line with the results in Table A1, which illustrates that intensive tracking was successful in re-establishing baseline balance. Nevertheless, attrition remains selective, as many baseline characteristics are significant predictors of attrition (Table 4).

[37]To proxy for temporary residence in the village, we consider whether the household owned the house of residence, specifically mentions whether the house was obtained in exchange for labor services (which is common in haciendas), has an address referring to the proximity or presence in an hacienda, or alternatively, a variable capturing a hacienda address and a house that is not owned by the household. Finally, Table 5 includes two additional proxies for locations with more temporary residents: the level of attrition between the census and the first baseline survey (i.e. between spring and fall of 2000) in the locality of origin of the individual, the share of individuals attrited and whether nobody attrited. The baseline survey was conducted right after the public lottery and before the start of the transfers. People coming from such locations not only were more likely to attrit, but also could be harder to trace back, as contacts with the community of origin can be limited.

Table B1: Correlates of Attrition - Complete Tracking Phase

	Early Treatment	Late Treatment	P-value Diff. ET and LT	P-value Joint Test
Variables Indicating Very Early Attrition				
Probability of attrition prior to program start in locality	-0.323	-0.211	0.781	0.235
	(0.375)	(0.140)		
Nobody of target sample attrited before program start	0.088**	0.093***	0.914	0.001
	(0.040)	(0.026)		
Proxy's of permanent residence in village				
Own house (= 1)	0.065	0.146**	0.381	0.054
	(0.065)	(0.064)		
House is obtained in exchange for service/labor (= 1)	-0.120	-0.154**	0.798	0.063
	(0.113)	(0.071)		
Address in hacienda (= 1)	-0.035	-0.111	0.446	0.277
	(0.068)	(0.072)		
Address in hacienda & house rented (= 1)	-0.059	-0.289**	0.153	0.091
	(0.082)	(0.136)		
Social Capital				
Family network size (individuals)	0.001***	0.001***	0.733	0.000
	(0.000)	(0.000)		
Population size village	-0.000	-0.000	0.948	0.844
	(0.000)	(0.000)		
Village Characteristics				
Village affected by hurrican Mitch (= 1)	-0.002	-0.096***	0.059	0.001
	(0.042)	(0.024)		
Altitude of village	-0.000	0.000	0.620	0.856
	(0.000)	(0.000)		
Village in coffee producing region (= 1)	-0.002	-0.004	0.964	0.992
	(0.048)	(0.036)		
Distance to night light (meters)	-0.000***	-0.000	0.105	0.000
	(0.000)	(0.000)		
Live in Tuma region (= 1)	-0.135***	-0.123***	0.764	0.000
	(0.028)	(0.032)		
Live in Madriz region (= 1)	0.128***	0.037	0.019	0.000
	(0.025)	(0.028)		
Household Characteristics: Economic activities* & *Assets				
Household head main occupation is ag. (= 1)	0.055	0.063	0.914	0.270
	(0.057)	(0.047)		
Size of landholdings ('000 sq meters)	-0.001	0.000	0.550	0.823
	(0.002)	(0.001)		
Number of parcels of land	0.093**	0.030	0.149	0.022
	(0.038)	(0.020)		
Log predicted expenditures (pc)	0.013	-0.121*	0.141	0.195
	(0.060)	(0.066)		
Wealth index - housing characteristics	-0.007	-0.024*	0.399	0.205
	(0.015)	(0.014)		
Wealth index - productive assets	0.034*	0.034**	0.981	0.018
	(0.017)	(0.015)		
Wealth index - other assets	-0.038***	-0.011	0.173	0.020
	(0.013)	(0.014)		

Table B1: Correlates of attrition - Complete Tracking Phase (Continue)

	Early Treatment	Late Treatment	P-value Diff. ET and LT	P-value Joint Test
Household Characteristics: Demographics				
Father not living in same household (= 1)	-0.116***	-0.099*	0.794	0.005
	(0.040)	(0.050)		
Mother not living in same household (= 1)	-0.061	-0.120	0.645	0.327
	(0.091)	(0.088)		
Child of household head (= 1)	0.045	0.057	0.890	0.494
	(0.067)	(0.057)		
Number of children of household head	0.006	0.019***	0.201	0.002
	(0.009)	(0.005)		
Female household head (= 1)	-0.111**	-0.046	0.365	0.062
	(0.049)	(0.051)		
Age of household head	0.002	0.002	0.930	0.324
	(0.002)	(0.002)		
Number of household members	-0.005	0.011	0.137	0.231
	(0.008)	(0.007)		
Nuclear household (= 1)	0.028	0.038	0.845	0.378
	(0.041)	(0.031)		
Multigenerational household (= 1)	0.077*	-0.029	0.052	0.132
	(0.041)	(0.033)		
Other household structure (= 1)	-0.159**	-0.035	0.200	0.080
	(0.071)	(0.063)		
Number of children aged 0-8	-0.040**	0.018	0.013	0.043
	(0.019)	(0.012)		
Number of children aged 9-12	-0.048	0.023	0.076	0.145
	(0.036)	(0.015)		
Household Characteristics: Education				
Distance to nearest school (minutes)	-0.000	-0.000	0.806	0.276
	(0.000)	(0.000)		
Mother no grades attained (= 1)	0.069	0.100**	0.629	0.024
	(0.050)	(0.040)		
Mother 3 plus grades attained (= 1)	-0.024	-0.008	0.816	0.888
	(0.063)	(0.027)		
Household head no grades attained (= 1)	0.003	-0.039	0.446	0.463
	(0.044)	(0.031)		
Household head 3 plus grades attained (= 1)	0.030	-0.006	0.488	0.722
	(0.038)	(0.035)		
Individual Characteristics				
Age at start of transfer in months	-0.041***	-0.024**	0.360	0.004
	(0.014)	(0.011)		
Highest grade attained	0.009	-0.010	0.477	0.642
	(0.023)	(0.012)		
No grades attained (= 1)	-0.068	-0.012	0.358	0.418
	(0.053)	(0.029)		
Worked in last week (= 1)	-0.046	0.005	0.377	0.597
	(0.045)	(0.035)		
Participated in some economic activity (= 1)	-0.026	-0.007	0.695	0.782
	(0.040)	(0.024)		
Obs.	588	550		

Note: Estimates for differences control for stratification fixed effects. Standard errors are clustered at locality level. Robust s.e. in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

B.2 Compliance

As with most CCT programs, compliance was high. Among the target population, 94 percent of boys in the early treatment group, and 90 percent of boys in the late treatment group lived in households that received at least one transfer.[38] We hypothesize that a lot of the non-compliance relates to migration out of the study area in the early years of the intervention. To the extent that these are also children that are hard to track, inclusion of these children in the sample will likely affect the ITT estimates.[39] It is therefore useful to investigate the correlates of compliance (Table B2), and compare them with the correlates of attrition.

For both experimental groups, compliance is significantly higher for households with large family networks in the village, and lower for more remote areas (as measured by distance to night light). Compliance for both groups is also lower for the Tuma region (a coffee producing area with large haciendas and temporary workers) while it is very high in Madriz (a region with high share of indigenous population). Not surprisingly, compliance is significantly lower in areas with very early migration (as measured by the attrition between the program census and baseline survey). Overall, these patterns are consistent with non-compliers being temporary residents in the intervention villages that had moved out before they could benefit from the program.

In the late treatment localities, households with an address in the hacienda are less likely to comply, while households with more parcels and more productive assets (animals), and individuals who were economically active at baseline are more likely to have complied. On the other hand, non-nuclear households and households without the father living in the households are less likely to have complied. These patterns are consistent

[38] We consider whether households received at least one transfer, rather than whether they received the totality of transfers they were entitled too. All households that registered for the program would have at least received one transfer, while receiving the totality of the transfers was also a function of compliance with conditionalities.

[39] The impact of the bias is hard to predict. Inclusion of non-compliers from the early treatment group should lower the ITT. But inclusion of the non-compliers from the late treatment group may have an ambiguous effect. It could not affect the estimates to the extent that these boys were too old for their education to benefit from the program. But it could also lower the estimates of their family benefiting would still have affected their education even if they were passed the eligibility age by 2003. As the non-compliance rate was slightly higher among the late treatment group, the overall bias could go either way.

with households with less economic or family ties migrating out of the study area prior to the start of the program in the late treatment group. We do not observe the same patterns in the early treatment group, though the difference between the groups is only significant for the nuclear household variable.

On the other hand, non-compliance in the early treatment group is significantly correlated with the boy having no education at baseline, while (somewhat contradictory) having a mother without education makes a household more likely to comply. Overall, there are few significant correlates for the early treatment group, consistent with the high compliance rate.

Table B2: Correlates of compliance

	Early Treatment	Late Treatment	P-value Diff. ET and LT	P-value Joint Test
Variables Indicating Very Early Attrition				
Probability of attrition prior to program start in locality	-0.460***	-0.369***	0.579	0.000
	(0.117)	(0.114)		
Nobody of target sample attrited before program start	0.064**	0.083**	0.662	0.004
	(0.026)	(0.032)		
Proxy's of permanent residence in village				
Own house (= 1)	0.037	0.101	0.415	0.211
	(0.034)	(0.070)		
House is obtained in exchange for service/labor (= 1)	-0.028	-0.119	0.532	0.510
	(0.035)	(0.140)		
Address in hacienda (= 1)	-0.037	-0.092**	0.363	0.090
	(0.038)	(0.045)		
Address in hacienda & house rented (= 1)	-0.056	-0.124	0.502	0.155
	(0.039)	(0.092)		
Social Capital				
Family network size (individuals)	0.000***	0.001***	0.293	0.000
	(0.000)	(0.000)		
Population size village	-0.000	-0.000	0.610	0.637
	(0.000)	(0.000)		
Village Characteristics				
Village affected by hurrican Mitch (= 1)	-0.044	0.028	0.202	0.223
	(0.026)	(0.049)		
Altitude of village	-0.000	0.000	0.274	0.467
	(0.000)	(0.000)		
Village in coffee producing region (= 1)	-0.042*	-0.009	0.381	0.182
	(0.023)	(0.029)		
Distance to night light (meters)	-0.000**	-0.000**	0.499	0.015
	(0.000)	(0.000)		
Live in Tuma region (= 1)	-0.076***	-0.096**	0.661	0.001
	(0.025)	(0.038)		
Live in Madriz region (= 1)	0.039*	0.091***	0.128	0.001
	(0.019)	(0.027)		
Household Characteristics: Economic activities & ***Assets***				
Household head main occupation is ag. (= 1)	-0.006	0.088*	0.106	0.237
	(0.024)	(0.051)		
Size of landholdings ('000 sq meters)	0.001**	0.000	0.626	0.077
	(0.000)	(0.001)		
Number of parcels of land	0.018	0.053*	0.273	0.070
	(0.013)	(0.028)		
Log predicted expenditures (pc)	-0.036	-0.028	0.919	0.621
	(0.041)	(0.064)		
Wealth index - housing characteristics	-0.008	-0.027	0.350	0.208
	(0.010)	(0.017)		
Wealth index - productive assets	0.018	0.040**	0.355	0.055
	(0.016)	(0.018)		
Wealth index - other assets	-0.006	-0.012	0.695	0.475
	(0.007)	(0.013)		

Table B2: Correlates of compliance (Continue)

	Early Treatment	Late Treatment	P-value Diff. ET and LT	P-value Joint Test
Household Characteristics: Demographics				
Father not living in same household (= 1)	-0.038	-0.094**	0.284	0.050
	(0.026)	(0.045)		
Mother not living in same household (= 1)	-0.022	-0.134	0.214	0.218
	(0.037)	(0.080)		
Child of household head (= 1)	0.039	0.117	0.330	0.149
	(0.034)	(0.072)		
Number of children of household head	0.006	0.012	0.580	0.189
	(0.005)	(0.009)		
Female household head (= 1)	-0.018	-0.041	0.717	0.623
	(0.026)	(0.059)		
Age of household head	0.000	-0.004	0.119	0.266
	(0.001)	(0.003)		
Number of household members	0.005	-0.003	0.320	0.281
	(0.003)	(0.007)		
Nuclear household (= 1)	-0.020	0.086**	0.006	0.023
	(0.017)	(0.033)		
Multigenerational household (= 1)	0.038	-0.058	0.083	0.146
	(0.024)	(0.049)		
Other household structure (= 1)	-0.017	-0.094	0.406	0.406
	(0.020)	(0.089)		
Number of children aged 0-8	0.003	-0.004	0.536	0.816
	(0.007)	(0.009)		
Number of children aged 9-12	-0.007	-0.041	0.317	0.354
	(0.016)	(0.029)		
Household Characteristics: Education				
Distance to nearest school (minutes)	-0.000	-0.001	0.440	0.282
	(0.000)	(0.001)		
Mother no grades attained (= 1)	0.030**	-0.014	0.088	0.030
	(0.011)	(0.023)		
Mother 3 plus grades attained (= 1)	-0.054**	0.013	0.074	0.093
	(0.025)	(0.027)		
Household head no grades attained (= 1)	0.024	0.030	0.860	0.202
	(0.015)	(0.032)		
Household head 3 plus grades attained (= 1)	-0.026	-0.059	0.513	0.223
	(0.022)	(0.046)		
Individual Characteristics				
Age at start of transfer in months	-0.008	0.015	0.192	0.412
	(0.011)	(0.013)		
Highest grade attained	0.009	0.001	0.614	0.606
	(0.008)	(0.013)		
No grades attained (= 1)	-0.042*	-0.015	0.551	0.176
	(0.023)	(0.038)		
Worked in last week (= 1)	-0.030	0.067**	0.100	0.058
	(0.050)	(0.028)		
Participated in some economic activity (= 1)	-0.009	0.072***	0.132	0.024
	(0.046)	(0.025)		
Obs.	588	550		

Note: Estimates for differences control for stratification fixed effects. Standard errors are clustered at locality level. Robust s.e. in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

B.3 Attrition after Regular Tracking Phase

For both experimental groups, the probability of being found in the regular phase is significantly higher for households with large family networks in the village, and lower for more remote areas (as measured by distance to light). The probability of being found in the regular phase is also significantly lower in areas with very early migration (as measured by the attrition between the program census and baseline survey) and the point estimate is large. These patterns reflect the findings for compliance.

The probability of being found in the regular phase is lower for the Tuma region and higher in Madriz, but only for the early treatment group. For the early treatment group, ownership of the house is a strong positive predictor, while those who rent a house on a hacienda address, or those that obtained a house for services are significantly less likely to be found. Households with more land or productive assets at baseline, or higher estimated per capita expenditures, and the household working in agriculture, are also more likely to be found, but the correlations are only significant for the early treatment group. Boys from households with more very young children, complex household structures, and large household sizes, in the early treatment were also less likely to be found.

On the other hand, being from a nuclear household, being the son of the household head, and the number of children of the household head, are all positively correlated to being found in the late treatment, while being from a female-head household, or a household where the father or the mother is not present is negatively correlated. These demographic characteristics appear to matter less for the early treatment group, and for most of these characteristics the differences between groups are significantly different. Education of the mother is negatively correlated with being found in the late treatment, and this is also significantly different than for the early treatment.

The different patterns between early and late treatment groups are striking and the demographic characteristics are consistent with the treatment having kept boys with weaker ties to the baseline households for longer in these households (consistent with the early presence of transfers). Boys in the late treatment group may have moved out of these households prior to 2003 or afterwards. In any case, the transfers would have provided less of an incentive for them to stay, as they aged out of the conditionalities. In contrast, the probability of being found in the early treatment group appears to be more related to the economic ties/opportunities.

Overall, the large number of significant covariates, and the significant differences between experimental groups clearly indicate that attrition is both selective and driven by different factors in the early versus late treatment group. The later explains the lack of balance in baseline characteristics after the regular tracking, with boys in the early treatment group in particular being less likely to be the son of the household head, coming from households with female headed households, and from mothers with higher levels of education. Estimates after the regular tracking hence likely would be biased. The selectivity of the attrition in both experimental groups moreover implies that the ITT estimates would not accurately reflect intent-to-treat estimates for the entire target population, in the presence of heterogeneous treatment effects. Given that some of the predictors of attrition are the same as the predictors of compliance, heterogeneous treatment effects moreover seem very likely.

Table B3: Correlates of attrition - Regular Tracking Phase

	Early Treatment	Late Treatment	P-value Diff. ET and LT	P-value Joint Test
Variables Indicating Very Early Attrition				
Probability of attrition prior to program start in locality	-0.716	-0.222	0.371	0.220
	(0.505)	(0.208)		
Nobody of target sample attrited before program start	0.164***	0.090**	0.319	0.005
	(0.060)	(0.042)		
Proxy's of permanent residence in village				
Own house (= 1)	0.213**	0.073	0.251	0.049
	(0.090)	(0.080)		
House is obtained in exchange for service/labor (= 1)	-0.433***	-0.104	0.024	0.000
	(0.085)	(0.112)		
Address in hacienda (= 1)	-0.207	-0.123	0.570	0.080
	(0.126)	(0.075)		
Address in hacienda & house rented (= 1)	-0.290*	-0.200	0.645	0.052
	(0.153)	(0.120)		
Social Capital				
Family network size (individuals)	0.001***	0.001***	0.920	0.000
	(0.000)	(0.000)		
Population size village	-0.000	0.000	0.639	0.891
	(0.000)	(0.000)		
Village Characteristics				
Village affected by hurrican Mitch (= 1)	0.025	-0.038	0.389	0.667
	(0.054)	(0.050)		
Altitude of village	-0.000	0.000	0.317	0.579
	(0.000)	(0.000)		
Village in coffee producing region (= 1)	-0.029	0.050	0.391	0.610
	(0.037)	(0.083)		
Distance to night light (meters)	-0.000***	-0.000*	0.027	0.000
	(0.000)	(0.000)		
Live in Tuma region (= 1)	-0.222***	-0.072	0.021	0.000
	(0.045)	(0.044)		
Live in Madriz region (= 1)	0.198***	0.007	0.029	0.003
	(0.053)	(0.065)		
Household Characteristics: Economic activities & ***Assets***				
Household head main occupation is ag. (= 1)	0.116*	0.072	0.641	0.095
	(0.058)	(0.074)		
Size of landholdings ('000 sq meters)	-0.000	0.001	0.294	0.338
	(0.001)	(0.001)		
Number of parcels of land	0.098*	0.062	0.618	0.092
	(0.057)	(0.042)		
Log predicted expenditures (pc)	0.137*	-0.127	0.023	0.068
	(0.074)	(0.083)		
Wealth index - housing characteristics	0.009	-0.035*	0.094	0.208
	(0.015)	(0.020)		
Wealth index - productive assets	0.058***	0.029	0.371	0.012
	(0.020)	(0.025)		
Wealth index - other assets	-0.068***	0.007	0.009	0.001
	(0.016)	(0.022)		

Table B3: Correlates of attrition - Regular Tracking Phase (Continue)

	Early Treatment	Late Treatment	P-value Diff. ET and LT	P-value Joint Test
Household Characteristics: Demographics				
Father not living in same household (= 1)	-0.090	-0.229***	0.067	0.000
	(0.057)	(0.046)		
Mother not living in same household (= 1)	-0.043	-0.285**	0.106	0.069
	(0.081)	(0.122)		
Child of household head (= 1)	-0.066	0.223***	0.002	0.001
	(0.068)	(0.055)		
Number of children of household head	-0.007	0.029***	0.003	0.003
	(0.008)	(0.008)		
Female household head (= 1)	-0.055	-0.151**	0.305	0.067
	(0.064)	(0.067)		
Age of household head	0.001	-0.001	0.666	0.910
	(0.002)	(0.003)		
Number of household members	-0.011	0.006	0.091	0.227
	(0.007)	(0.007)		
Nuclear household (= 1)	0.068	0.104**	0.555	0.021
	(0.042)	(0.043)		
Multigenerational household (= 1)	0.029	-0.114**	0.034	0.096
	(0.035)	(0.055)		
Other household structure (= 1)	-0.165**	-0.018	0.216	0.099
	(0.075)	(0.090)		
Number of children aged 0-8	-0.040***	-0.005	0.041	0.001
	(0.010)	(0.013)		
Number of children aged 9-12	-0.054	0.010	0.180	0.375
	(0.040)	(0.025)		
Household Characteristics: Education				
Distance to nearest school (minutes)	0.000	-0.001*	0.158	0.229
	(0.001)	(0.001)		
Mother no grades attained (= 1)	-0.059*	0.114***	0.002	0.006
	(0.034)	(0.038)		
Mother 3 plus grades attained (= 1)	0.073	-0.070*	0.051	0.114
	(0.059)	(0.040)		
Household head no grades attained (= 1)	-0.040	-0.027	0.843	0.582
	(0.046)	(0.046)		
Household head 3 plus grades attained (= 1)	0.073	-0.006	0.312	0.507
	(0.062)	(0.046)		
Individual Characteristics				
Age at start of transfer in months	-0.041**	-0.016	0.391	0.060
	(0.017)	(0.022)		
Highest grade attained	0.017	-0.019	0.210	0.385
	(0.023)	(0.016)		
No grades attained (= 1)	-0.081*	0.013	0.185	0.224
	(0.046)	(0.052)		
Worked in last week (= 1)	-0.065	-0.023	0.567	0.453
	(0.055)	(0.046)		
Participated in some economic activity (= 1)	-0.041	-0.019	0.755	0.678
	(0.052)	(0.045)		
Obs.	588	550		

Note: Estimates for differences control for stratification fixed effects. Standard errors are clustered at locality level. Robust s.e. in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table B4: Correlates of attrition - Intensive Tracking Phase

	Early Treatment	Late Treatment	P-value Diff. ET and LT	P-value Joint Test
Variables Indicating Very Early Attrition				
Probability of attrition prior to program start in locality	-0.119	-0.534	0.588	0.471
	(0.623)	(0.437)		
Nobody of target sample attrited before program start	0.081	0.253**	0.184	0.025
	(0.085)	(0.094)		
Proxy's of permanent residence in village				
Own house (= 1)	-0.052	0.389***	0.011	0.000
	(0.145)	(0.081)		
House is obtained in exchange for service/labor (= 1)	0.091	-0.334***	0.041	0.005
	(0.176)	(0.098)		
Address in hacienda (= 1)	0.113	-0.184	0.129	0.235
	(0.152)	(0.118)		
Address in hacienda & house rented (= 1)	0.114	-0.494***	0.010	0.001
	(0.185)	(0.127)		
Social Capital				
Family network size (individuals)	0.002***	0.002**	0.852	0.001
	(0.001)	(0.001)		
Population size village	-0.000	-0.000	0.583	0.832
	(0.000)	(0.000)		
Village Characteristics				
Village affected by hurrican Mitch (= 1)	-0.041	-0.374***	0.034	0.002
	(0.114)	(0.100)		
Altitude of village	-0.000	-0.000	0.970	0.882
	(0.000)	(0.000)		
Village in coffee producing region (= 1)	0.037	-0.092	0.475	0.667
	(0.144)	(0.106)		
Distance to night light (meters)	-0.000***	-0.000	0.419	0.008
	(0.000)	(0.000)		
Live in Tuma region (= 1)	-0.192***	-0.348***	0.202	0.000
	(0.068)	(0.100)		
Live in Madriz region (= 1)	0.323***	0.143*	0.056	0.000
	(0.048)	(0.078)		
Household Characteristics: Economic activities* & *Assets				
Household head main occupation is ag. (= 1)	0.022	0.113	0.630	0.440
	(0.165)	(0.088)		
Size of landholdings ('000 sq meters)	-0.002	-0.002	0.955	0.434
	(0.003)	(0.002)		
Number of parcels of land	0.142*	0.018	0.216	0.210
	(0.080)	(0.057)		
Log predicted expenditures (pc)	-0.127	-0.185	0.774	0.311
	(0.130)	(0.154)		
Wealth index - housing characteristics	-0.035	-0.027	0.870	0.395
	(0.031)	(0.033)		
Wealth index - productive assets	0.049	0.089***	0.520	0.009
	(0.054)	(0.028)		
Wealth index - other assets	-0.045	-0.058	0.830	0.212
	(0.044)	(0.040)		

Table B4: Correlates of attrition - Intensive Tracking Phase (Continue)

	Early Treatment	Late Treatment	P-value Diff. ET and LT	P-value Joint Test
Household Characteristics: Demographics				
Father not living in same household (= 1)	-0.234***	-0.009	0.143	0.031
	(0.085)	(0.124)		
Mother not living in same household (= 1)	-0.134	-0.001	0.609	0.816
	(0.209)	(0.150)		
Child of household head (= 1)	0.292**	-0.094	0.050	0.103
	(0.141)	(0.129)		
Number of children of household head	0.026	0.030	0.894	0.188
	(0.021)	(0.022)		
Female household head (= 1)	-0.265**	0.048	0.052	0.034
	(0.099)	(0.122)		
Age of household head	0.006	0.007*	0.910	0.068
	(0.005)	(0.004)		
Number of household members	-0.003	0.035	0.178	0.286
	(0.017)	(0.022)		
Nuclear household (= 1)	0.002	-0.021	0.874	0.978
	(0.104)	(0.102)		
Multigenerational household (= 1)	0.234**	0.067	0.266	0.108
	(0.114)	(0.094)		
Other household structure (= 1)	-0.232**	-0.107	0.418	0.061
	(0.103)	(0.114)		
Number of children aged 0-8	-0.061	0.079**	0.024	0.044
	(0.049)	(0.035)		
Number of children aged 9-12	-0.074	0.121	0.101	0.252
	(0.066)	(0.095)		
Household Characteristics: Education				
Distance to nearest school (minutes)	-0.001	0.000	0.223	0.361
	(0.001)	(0.001)		
Mother no grades attained (= 1)	0.311**	0.232	0.697	0.031
	(0.139)	(0.144)		
Mother 3 plus grades attained (= 1)	-0.200	0.078	0.143	0.336
	(0.154)	(0.105)		
Household head no grades attained (= 1)	0.065	-0.112	0.293	0.519
	(0.126)	(0.109)		
Household head 3 plus grades attained (= 1)	0.002	-0.014	0.913	0.992
	(0.088)	(0.106)		
Individual Characteristics				
Age at start of transfer in months	-0.084***	-0.071*	0.778	0.005
	(0.029)	(0.037)		
Highest grade attained	0.005	-0.009	0.845	0.975
	(0.057)	(0.042)		
No grades attained (= 1)	-0.118	-0.070	0.804	0.629
	(0.152)	(0.121)		
Worked in last week (= 1)	-0.058	0.058	0.446	0.742
	(0.100)	(0.114)		
Participated in some economic activity (= 1)	-0.029	0.003	0.788	0.945
	(0.088)	(0.083)		
Obs.	160	137		

Note: Estimates for differences control for stratification fixed effects.Standard errors are clustered at locality level. Robust s.e. in parentheses.*** p<0.01, ** p<0.05, * p<0.1.

B.4 Linear Probability Model. Relation of covariates used to compute the IPW.

Table B5: Linear Probability Model - Correlates of attrition

	CTP Sample		RTP Sample		ITP Sample	
		ET Interaction		ET Interaction		ET Interaction
	(1)	(2)	(3)	(4)	(5)	(6)
Early-Treatment	0.691		0.103		-0.581	
	(0.82)		(0.29)		(0.38)	
Very early attrition (Z)						
Probability of attrition prior to program start in locality			-0.002	-0.775***		
			(0.15)	(0.28)		
Nobody of target sample attrited before program strat	0.047**	0.004			0.120	-0.075
	(0.02)	(0.03)			(0.11)	(0.15)
Proxy's of permanent residence in village (Z)						
Own house (= 1)					0.371**	-0.331*
					(0.15)	(0.18)
House is obtained in exchange for service/labor (= 1)	0.060	-0.090	0.028	-0.335**	0.407***	-0.356*
	(0.10)	(0.14)	(0.11)	(0.13)	(0.12)	(0.21)
Address in hacienda & house rented (= 1)	-0.205	0.232			-0.231	0.427
	(0.14)	(0.16)			(0.23)	(0.27)
Networks (Z)						
Family network size (individuals)	0.000**	0.000	0.001***	-0.000	0.001	0.001
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Village Characteristics (Z)						
Village affected by hurricane Mitch (= 1)	-0.105***	0.111**			-0.278**	0.163
	(0.03)	(0.05)			(0.12)	(0.15)
Distance to night light (meters)	0.000	-0.000**	-0.000	-0.000**		
	(0.00)	(0.00)	(0.00)	(0.00)		
Live in Tuma region (= 1)	-0.079***	0.151***	0.039	0.108	-0.283***	0.379***
	(0.03)	(0.05)	(0.04)	(0.08)	(0.08)	(0.11)
Live in Madriz region (= 1)	0.097***	-0.046	0.026	0.007	0.179	-0.079
	(0.04)	(0.05)	(0.06)	(0.08)	(0.13)	(0.19)
Household Characteristics: Economic activities* & *Assets (X)						
Household head main occupation is ag. (= 1)			-0.001	0.101		
			(0.05)	(0.09)		
Log of size of landholdings	-0.004	0.018**	0.005	0.015		
	(0.00)	(0.01)	(0.01)	(0.01)		
Number of parcels of land			0.027	-0.108	-0.168**	0.303***
			(0.04)	(0.08)	(0.07)	(0.09)
Log predicted expenditures (pc)	0.002	-0.110				
	(0.05)	(0.11)				
Wealth index - housing characteristics	-0.014	0.038*	-0.022	0.057**		
	(0.01)	(0.02)	(0.02)	(0.02)		
Wealth index - productive assets	0.036***	-0.022	0.018	-0.010	0.058	-0.033
	(0.01)	(0.02)	(0.02)	(0.03)	(0.03)	(0.06)

Table B5: Linear Probability Model - Correlates of attrition (Continue)

	CTP Sample		RTP Sample		ITP Sample	
		ET Interaction		ET Interaction		ET Interaction
	(1)	(2)	(3)	(4)	(5)	(6)
Household Characteristics: Demographics (X,Z)						
Father not living in same household (= 1)	-0.086*	0.054	-0.157*	0.106	-0.144	0.134
	(0.04)	(0.06)	(0.08)	(0.12)	(0.17)	(0.19)
Mother not living in same household (= 1)			-0.124	-0.043		
			(0.11)	(0.13)		
Child of household head (= 1)			0.005	-0.137	0.050	0.209
			(0.09)	(0.12)	(0.18)	(0.22)
Number of children of household head			0.012	-0.015		
			(0.01)	(0.02)		
Female household head (= 1)	0.073	-0.098	0.089	-0.039	0.374***	-0.506***
	(0.06)	(0.08)	(0.08)	(0.12)	(0.12)	(0.18)
Nuclear household (= 1)			-0.048	0.227***		
			(0.05)	(0.08)		
Multigenerational household (= 1)			-0.088	0.281***		
			(0.06)	(0.09)		
Other household structure (= 1)	0.037	-0.173***			0.426***	-0.434***
	(0.03)	(0.06)			(0.12)	(0.15)
Number of children aged 0-8	0.017	-0.050***	-0.019	-0.004	0.072***	-0.131***
	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.04)
Household Characteristics: Education (X)						
Distance to nearest school (minutes)			-0.001*	0.003***		
			(0.00)	(0.00)		
Mother no grades attained (= 1)	0.120***	-0.059	0.102**	-0.116**	0.159*	0.022
	(0.03)	(0.06)	(0.04)	(0.06)	(0.09)	(0.13)
Stratification fixed effects (X)	YES	YES	YES	YES	YES	YES
Three monthly age dummies	YES	YES	YES	YES	YES	YES
RTP survey supervisor fixed effects	YES	YES	YES	YES	YES	YES
Obs.	1138	1138	1138	1138	297	297
R-squared	0.23	0.23	0.22	0.22	0.51	0.51
adjusted R-squared	0.19	0.19	0.16	0.16	0.38	0.38
Joint Significance Tests by group of covariates						
Covariates and interaction terms together						
F-stat region, residence & network (Z)	12.131		6.145		11.186	
F-stat demogr. (X,Z)	3.812		4.708		3.671	
F-stat AGE	4.726		3.017		4.291	
F-stat SES & STRATA (X)	8.021		9.093		5.439	
Covariates and interaction terms separately						
F-stat region, residence & network (Z)	16.216	4.242	2.410	4.131	16.032	11.551
F-stat demogr. (X,Z)	2.593	3.837	3.464	2.328	4.808	6.687
F-stat AGE	6.364	1.685	3.090	1.518	3.703	2.871
F-stat SES & STRATA (X)	6.194	3.319	5.348	4.106	5.292	4.378

Note: Each two columns report the results for one single equation, even columns showing the value of the coefficients on the interaction between variables listed in the first column and assignment to early treatment. The first four columns show the results on the probability to be found after conducting the CTP and the RTP, the last two columns show the results for the selection model on the sample of individuals targeted during the ITP. Standard errors are clustered at locality level. Robust s.e. in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

C Appendix: Field protocols to track migrant households and individuals in RPS long-term follow-up survey in Nicaragua

Extensive tracking procedures were a key component of the research design for the long-term evaluation of RPS: all households (of the original women beneficiary), as well as all individuals under age 22 were targeted for follow-up. All households and individuals that could not be found in their original locations were tracked to their new locations, wherever they went in Nicaragua. Migrants to Costa Rica (the destination of over 95 percent of international migrants from the sample) also were tracked. As migration is often temporary, multiple return visits to the original locations by the field team were organized to incorporate temporary migrants who might have returned. This allowed either to find the migrant directly, in case they had returned to the home location, or to update the contact and destination information.

C.1 Information collected about migrants

All households and individuals were initially sought using the direction of their home registered in the baseline program census. Survey teams always had a list of names of all original household members, their age, national identifier number, relationships and other key characteristics – to facilitate the search and assure correct identification of targeted households and individuals. Using this information, survey teams consulted community leaders and other community members to locate households. Each time a targeted individual or household could not be found, information regarding their potential destination, contact information and other whereabouts was collected from three different sources, if at all possible: 1) the parents or another former household member (in case at least one of them was still in the community); 2) the leader of the community; and 3) another person from the community considered a friend, extended family member or a neighbor. For each migrant, the probable destination (municipality, village, direction) was recorded, the name and contact information of the head of the household where he/she could be found, as well as the name and contact of another person at the destination location. The later person could be, for instance, an employer or a family member that is possibly easier to locate – and once located could provide the information about the target person. Information was cross-validated between different sources. In case of doubt between different destinations, the team pursued search efforts in both destinations. The best-informed former household member would also be asked some key outcome questions about each migrant: in particular the years of education attained, their civil status, and their occupation. This provides proxy information on some of the key outcomes, in case the migrant ultimately cannot be tracked. When phone numbers were available, the teams tried to reach the migrant by phone, together with the person providing the contact information, both to validate the phone number and to increase the level of trust by the migrant. In those cases when former household members did not have phone contact information, they were asked to inquire about such contact information next time the migrant visited or called. This information was then recuperated by the survey team through follow-ups by phone or through new visits. Finally, during search efforts in new destinations, each migrant that was successfully located would be asked for information about the other migrants of his original community (or municipality) thought to be in the same destination. In addition, former household members that might have migrated to different locations, were asked contact information about each other.

C.2 Different phases of tracking

The target household sample consisted of 3,521 households from 12 municipalities, 21 percent of whom could not be interviewed in their original location. As a result of tracking efforts, final attrition at the household level is below 8 percent. This sample includes 757 households from a non-experimental comparison group in 6 municipalities neighboring the 6 RPS municipalities. These are not used in this paper. There is no significant difference in attrition between the experimental early and late treatment groups. At the individual level, of the 10,977 individuals under 22 years of age who were targeted to be included in the sample for the individual survey, 41 percent could not be interviewed in their original location when the survey team first visited. Of those, approximately 19 percent were temporarily absent, while the remaining 22 percent had migrated to other households, often in other locations in Nicaragua or Costa Rica. As a result of extensive tracking, final individual attrition due to permanent migration for those under 22 years old is 9 percent. For 5 percent, however, we have proxy information on the individual from the household survey. As with household-level attrition, there are no significant differences in attrition among early treatment, late treatment, and non-experimental comparison groups for individual attrition in the household survey, nor for attrition in the individual survey.[40]

These rates of attrition compare favorably to other impact evaluation or longitudinal studies covering similar or shorter periods and focusing on similar populations. In contrast to other longer-term studies, we tracked all households and targeted individuals, rather than a random subsample, to both increase statistical power and better capture heterogeneity that might be related to different destinations. Tracking was done in four stages.[41] First, during the regular survey period, carried out in all 12 original municipalities, survey teams tracked and immediately included all households and individuals that had moved to another location in the same or a nearby community. In cases of temporary absence, the team revisited the households in subsequent days while they were still working in the area. Location information for everyone else was recorded at this time. This regular survey period lasted approximately four months. Second, after completing this first round, tracking teams returned to each of the 12 study municipalities, finding and surveying many of the individuals and households who had been away during the earlier visit (temporary migrants). During this second round, at least two different teams visited each of the original localities. Households and individuals who had moved to any other locations within the original 12 municipalities also were tracked. Information collected previously on the destination of migrants was verified and updated. Concurrently,

[40] Tracking results and balance for the cohort used in this paper is discussed in section 3.4.

[41] Figures A2 and A3 in Appendix C provides information of the location of all permanent migrants who were found and successfully interviewed.

a separate team was tracking in Managua and its surroundings (the dominant migrant destination) and teams communicated in real time to update information on migrant movements. This second round lasted approximately four months. In the third round of tracking, we extended the search to all destinations in Nicaragua, both urban and rural; it lasted approximately seven months, including a three-month cessation of all field activities during the rainy (and hurricane) season when road access in many rural areas was very difficult and potentially dangerous. During this phase, field teams were simultaneously operating in different parts of the country and continued communicating all updated information on migrant movements. Lastly, additional return visits to original locations were undertaken during the Christmas-New Year break, as many migrants (temporarily) return to visit their family at that time of the year. In the last round, tracking was expanded to Costa Rica, the destination of the vast majority of international migrants from the sample. This final round lasted approximately two months.

C.3 Location of all permanent migrants

Figure A2: Location of permanent migrants found during the Regular Tracking Phase.

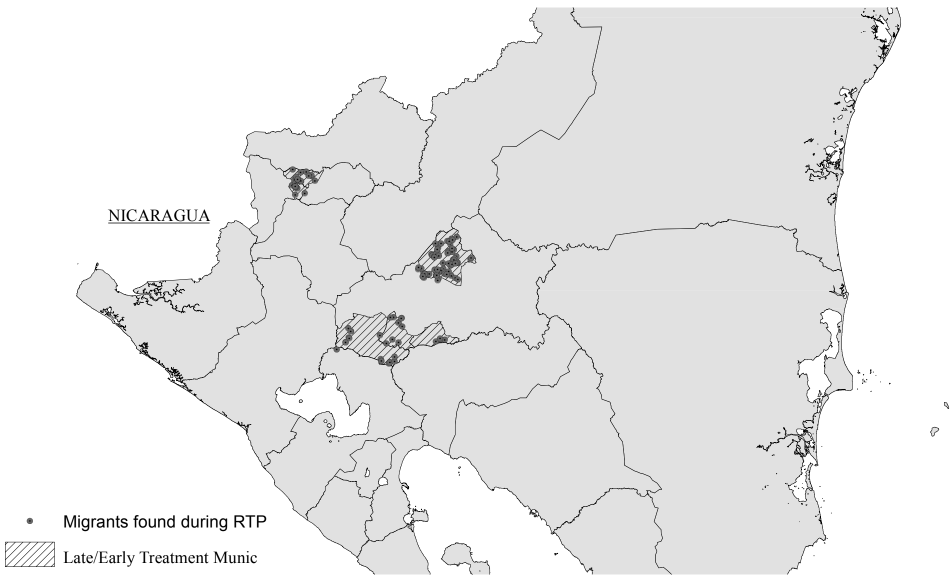
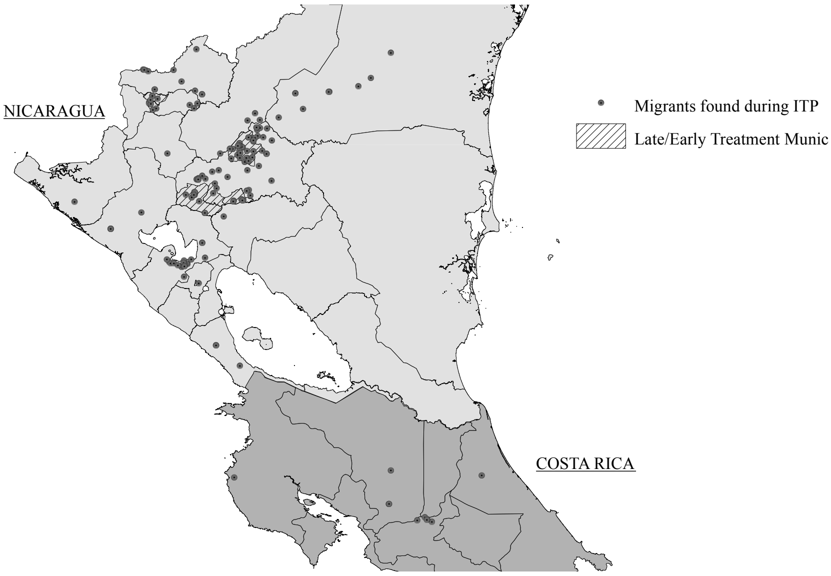


Figure A3: Location of permanent migrants found during the Intensive Tracking Phase.



D Appendix: Wealth index – principal component

The baseline program census contains a number of variables to proxy household wealth, including variables capturing characteristics of the house and assets owned. Following Filmer and Pritchett (2001) we aggregate these characteristics using principal component analysis. We do not include ownership of agricultural land or the house as these are likely to affect migration decisions directly. The principal component estimate is done using the baseline target sample. We retain the first three principal components, which jointly explain 53 percent of the variation, as they have an eigenvalue of more than 1. The first principal component mostly captures the characteristics of the house, while the second principal component has high weights for productive assets (i.e. ownership of animals and a fumigator), the third has high weights on specific amenities (zinc roof and latrine).

Table D1: Principal component scoring coefficients.

Variable	PC1	PC2	PC3
Work animals (= 1)	0.1312	0.6159	-0.0321
Fumigation sprayer (= 1)	0.1199	0.589	0.3619
Number of rooms in house	0.3489	0.2665	0.0138
Radio (= 1)	0.3851	0.068	-0.2142
Cement block walls (= 1)	0.4423	-0.1043	0.1099
Zinc roof (= 1)	0.2137	-0.1648	0.6923
Dirt floor (= 1)	-0.4367	0.2591	-0.1836
Latrine or toilet (= 1)	0.2805	0.144	-0.4941
Electric light (= 1)	0.4333	-0.2688	-0.2298

E Appendix: Discrete control variables

Table E1: Intent-To-Treat

	Complete tracking phase				Regular tracking phase				Intensive tracking phase			
	Stratif. Dummies	+Age&Educ. Controls	+Unbalanced Controls	+Baseline Controls	Stratif. Dummies	+Age& Educ. Controls	+Unbalanced Controls	+Baseline Controls	Stratif. Dummies	+Age & Educ . Controls	+Unbalanced Controls	+Baseline Controls
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Grades attained												
Early-Treatment	0.593**	0.427**	0.282*	0.248	0.865**	0.613***	0.412**	0.302	-0.745	-0.481	-0.527	-0.091
	(0.29)	(0.17)	(0.16)	(0.16)	(0.33)	(0.20)	(0.20)	(0.18)	(0.58)	(0.37)	(0.43)	(0.44)
Outcome Mean	5.45	5.45	5.45	5.45	5.39	5.39	5.39	5.39	5.81	5.81	5.81	5.81
R square	0.04	0.42	0.43	0.44	0.05	0.43	0.46	0.48	0.10	0.48	0.49	0.55
Comparing coefficients at different stages during the tracking process: P-values												
$ET^{CTP} - ET^{RTP} = 0$					0.0033	0.0361	0.1293	0.5134				
$ET^{CTP} - ET^{ITP} = 0$									0.0241	0.0149	0.0581	0.3875
$ET^{RTP} - ET^{ITP} = 0$									0.0231	0.0143	0.0580	0.3943
Off-Farm Employment												
Early-Treatment	0.059*	0.061*	0.056*	0.053**	0.087**	0.088**	0.087**	0.083***	-0.071	-0.097*	-0.147**	-0.197***
	(0.03)	(0.03)	(0.03)	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)	(0.05)	(0.06)	(0.07)	(0.06)
Outcome Mean	0.83	0.83	0.83	0.83	0.81	0.81	0.81	0.81	0.92	0.92	0.92	0.92
R square	0.02	0.04	0.06	0.08	0.03	0.05	0.07	0.10	0.04	0.24	0.29	0.44
Comparing coefficients at different stages during the tracking process: P-values												
$ET^{CTP} - ET^{RTP} = 0$					0.0041	0.0076	0.0051	0.0070				
$ET^{CTP} - ET^{ITP} = 0$									0.0093	0.0030	0.00151	0.000
$ET^{RTP} - ET^{ITP} = 0$									0.0073	0.0027	0.0012	0.000
Obs.	1006	1006	1006	1006	827	827	827	827	179	179	179	179

Note: The 1st model includes only strata fixed effects; in the 2nd model we add 3 monthly age fixed effects and set of dummies indicating whether individual had 1,2,3 or at least 4 years of education at baseline; additionally the 3rd model includes a vector of covariates that ended up off-balance after each of the tracking phases. After complete tracking the off-balance baseline controls are whether the individual was working, the number of individuals with family ties in village, the village population size and a productive asset index (2nd principal component of household wealth). After regular tracking the off-balance controls are mother with no education, mother with at least three years of education, the individual is son of the household head, the number of children of the household head and female household head. The regression on the sample targeted during ITP the off-balance baseline are the number of individuals with family ties in village, the village population size and a productive asset index (2nd principal component of household wealth); the 4th and last specification also controls for distance to school, number of children 0-8 and 9-12 in the household, estimated per capita consumption and estimated per capita consumption squared, as well as regional fixed effects. As in model 3, all categorical and continuous covariates are replaced with binary variables indicating whether individual is above the sample median for each of these variables. Standard errors are clustered at locality level. Robust s.e. in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table E2: Weighted Least Squares-Correcting for sample selection

	New IPW				Standard IPW			
	Stratif. Dummies	+Age&Educ. Controls	+Unbalanced Controls	+Baseline Controls	Stratif. Dummies	+Age& Educ. Controls	+Unbalanced Controls	+Baseline Controls
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Grades attained								
Complete Tracking Phase Sample								
Early-Treatment	0.361	0.363*	0.243	0.232	0.503	0.392**	0.261	0.229
	(0.32)	(0.18)	(0.17)	(0.16)	(0.30)	(0.19)	(0.18)	(0.16)
R-squared	0.03	0.42	0.43	0.45	0.04	0.41	0.42	0.44
Obs.	1006	1006	1006	1006	1006	1006	1006	1006
Regular Tracking Phase Sample								
Early-Treatment					0.770**	0.655***	0.501**	0.361**
					(0.32)	(0.21)	(0.20)	(0.18)
R-squared					0.04	0.43	0.46	0.48
Obs.					826	826	826	826
Off-Farm Employment								
Complete Tracking Phase Sample								
Early-Treatment	0.059**	0.061*	0.054*	0.051**	0.060**	0.063**	0.058*	0.054**
	(0.03)	(0.03)	(0.03)	(0.02)	(0.03)	(0.03)	(0.03)	(0.02)
R-squared	0.02	0.05	0.07	0.09	0.02	0.04	0.06	0.09
Obs.	1006	1006	1006	1006	1006	1006	1006	1006
Regular Tracking Phase Sample								
Early-Treatment					0.088**	0.086**	0.084**	0.081***
					(0.03)	(0.04)	(0.03)	(0.03)
R-squared					0.03	0.05	0.07	0.09
Obs.					827	827	827	827

Note: The 1st model includes only strata fixed effects; in the 2nd model we add 3 monthly age fixed effects and set of dummies indicating whether individual had 1,2,3 or at least 4 years of education at baseline; additionally the 3rd model includes a vector of covariates that ended up off-balance after each of the tracking phases. After complete tracking the off-balance baseline controls are whether the individual was working, the number of individuals with family ties in village, the village population size and a productive asset index (2nd principal component of household wealth). After regular tracking the off-balance controls are mother with no education, mother with at least three years of education, the individual is son of the household head, the number of children of the household head and female household head. The regression on the sample targeted during ITP the off-balance baseline are the number of individuals with family ties in village, the village population size and a productive asset index (2nd principal component of household wealth); the 4th and last specification also controls for distance to school, number of children 0-8 and 9-12 in the household, estimated per capita consumption and estimated per capita consumption squared, as well as regional fixed effects. As in model 3, all categorical and continuous covariates are replaced with binary variables indicating whether individual is above the sample median for each of these variables. Standard errors are clustered at locality level. Robust s.e. in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

F Appendix: LASSO methods for selection of predictors

As the potential set of predictors for attrition is large, section 5.2 uses bivariate regressions and stepwise selection of variables maximizing adjusted R-squared to obtain for each of the samples prediction models with the best possible predictions, following Doyle et al. (2016). This stepwise approach could raise, however, concerns of overfitting. We therefore obtain alternative prediction models using various LASSO procedures Tibshirani (1996). Following Ahrens, Hansen and Schaffer (2020), penalty levels for the LASSO were determined using cross-validation methods or information criteria, in particular the bias-corrected Akaike Information Criteria (AICc), developed for small samples, and shown to have good selection performance. We use the different LASSO models to obtain predictions for each of the treatment groups separately, and then use any covariates that were selected for prediction of attrition either in the treatment group or in the control group together with the interaction effect of those covariates and the treatment indicators. This is similar to the approach used for the stepwise selection and most consistent with the conceptual model that explicitly models differences in the relationship between covariates and attrition between treatment groups. It also has parallels with the post-double-selection methodology of Belloni, Chernozhukov and Hansen (2014). Alternatively, we use only one prediction for both treatment groups together, including in the list of possible regressors all covariates and their interaction with the treatment indicator. This results (logically) in a final prediction model that only includes the interaction with the treatment indicator for some regressors, but not the variable itself (and vice versa). It has the advantage, however, that it does not amplify the number of control variables, further addressing potential concerns of overfitting.

Table F1 shows results with the AICc estimators using these different alternatives. All results are broadly in line with results in Table 6, and notably confirm that estimates with new IPW are relatively more stable to inclusion of controls than estimates with the full IPW. The point estimates for the specification with controls are moreover similar in the different specifications, confirming overall robustness of results. All weights obtained are in the range of the original sample weights. Final models for both AICc approaches have similar R-squared as those presented in Table B5 for the full (R-squared =0.24-0.26), RTP (R-squared=0.21), and ITP (R-squared =0.53-0.54), and the models with the double prediction also end up with a similar number of covariates as those in Table B5 (17 variables and their interaction effects, compared to 16 variables and interactions in Table B5). These results together suggest that overfitting is not driving the results in Table 6, and more generally show robustness.

Using cross validation LASSO, however, leads to models that explain substantially less variation (R-squared =0.44) and more problematically leads to some very high weights on individual observations, which in turn leads to large standard errors in the IPW estimates and overall insignificant treatment effects across the different specifications.

Table F1: Weighted Least Squares-Correcting for sample selection

	New IPW				Standard IPW			
	Stratif. Dummies	+Age&Educ. Controls	+Unbalanced Controls	+Baseline Controls	Stratif. Dummies	+Age& Educ. Controls	+Unbalanced Controls	+Baseline Controls
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Grades attained								
AICc-double prediction model								
Early-Treatment	0.443	0.400**	0.363*	0.297*	0.508*	0.420**	0.360*	0.313*
	(0.30)	(0.19)	(0.19)	(0.17)	(0.30)	(0.19)	(0.18)	(0.16)
R-squared	0.04	0.41	0.42	0.45	0.05	0.44	0.45	0.49
AICc-combined prediction model								
Early-Treatment	0.359	0.426**	0.394*	0.330*	0.656**	0.464**	0.370**	0.331**
	(0.31)	(0.19)	(0.20)	(0.17)	(0.32)	(0.18)	(0.17)	(0.16)
R-squared	0.03	0.41	0.42	0.45	0.04	0.43	0.43	0.47
Off-Farm Employment								
AICc-double prediction model								
Early-Treatment	0.0709**	0.0674**	0.0596*	0.0547**	0.0601**	0.0617**	0.0571*	0.0545**
	(0.032)	(0.032)	(0.034)	(0.024)	(0.028)	(0.030)	(0.031)	(0.022)
R-squared	0.02	0.06	0.08	0.11	0.02	0.04	0.06	0.09
AICc-combined prediction model								
Early-Treatment	0.0563*	0.0565*	0.0519	0.0501**	0.0570*	0.0598*	0.0538	0.0538**
	(0.029)	(0.031)	(0.031)	(0.023)	(0.030)	(0.032)	(0.033)	(0.023)
R-squared	0.01	0.04	0.06	0.09	0.02	0.04	0.06	0.09

Note: N=1006. See Table 6 for details on specifications, and text in Appendix F for explanation on different prediction models. Standard errors are clustered at locality level. Robust s.e. in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

G Appendix: Literature Review

G.1 Strategies for dealing with attrition

Dealing with attrition *ex-ante* and *during the tracking process*

In contexts with high levels of mobility, as often found in developing countries, tracking migrants to their new destinations is often key to minimize attrition rates. While tracking migrants can imply high costs in terms of resources and time, experience from a number of high-quality non-experimental longitudinal studies show it to be feasible. The *Indonesia Family Life Survey* (IFLS) track households and selected household members from 1993 to 2014/15 within the 13 IFLS provinces. After 21 years the annual attrition rate for all target respondents who were in IFLS1 (1993) is less than 1 percent (accumulated attrition rate of 13 percent). See Tables G7-G8 in Appendix G.2 for a comparison of these attrition rates with other longitudinal studies that do not track nationally or internationally.[42] In the *Kagera Health and Development Survey* (KHDS) respondents were tracked within Tanzania and Uganda and the attrition rate is 12 percent of the panel survivors between 1994 and 2010. And in the second and third round of the *Mexican Family Life Survey* (MxFLS) (2002, 2005-2006 and 2009-2013) movers to the U.S. were tracked and interviewed in the U.S. and the accumulated attrition rate in the third follow-up is 13 percent.

Such attrition rates compare favorable with those of many RCT studies. RCT studies often do not include much information on the tracking protocols, suggesting it may be limited. Important exceptions, however, are a number of long-term follow-up studies that use intensive tracking, often only on a random subsample of those not found at the location of origin. This design was implemented in the 2002 follow-up survey of the

[42] Among the twenty six longitudinal databases reviewed, 58 percent were not designed to follow respondents beyond the borders of the village and many suffer from high attrition rate. The other surveys build various strategies for tracking beyond village borders. The common rule is to track individuals within the sample region (e.g. IFLS, Thomas et al. (2012)) or to popular migrants destination (e.g. the Kwazulu-Natal Income Dynamics Study, Alderman et al. (2001)). Only four surveys track individuals to any location within national borders and in three cases the tracking protocol includes following up migrants to other countries (Kagera Health and Development Survey, Mexican Family Life Survey and Albania Panel Survey).

U.S Moving to Opportunity (MTO) program (Orr et al., 2003) and has also been used, for instance, for impact evaluations surveys in developing countries (Kremer, Miguel and Thornton (2009); Blattman, Fiala and Martinez (2014); Duflo, Dupas and Kremer (2015); Baird et al. (2016); Blattman, Emeriau and Fiala (2018)). DiNardo, McCrary and Sanbonmatsu (2006) showed that in large enough samples tracking a random subsample of those missing to all their possible destinations provides a representative sample of the initial target population and estimates with high internal validity. Sample representativeness may however be hard to achieve with this approach when samples are small and the decision to migrate or the treatment estimates are heterogeneous.[43]

A notable study using intensive tracking on the full sample is, Duflo, Dupas and Kremer (2017), reaching response rates of about 98 percent after 8 years in Ghana by distributing mobile phones to every individual in the target population. This approach demonstrates that continuous efforts on keeping respondents' contact information updated (annual updates of contact information by phone or in person) and the technology activated (twice a year each member of the target population received mobile phone credit) can lead to very low attrition rates.

A third alternative to avoid high attrition rates ex-ante is to collect proxy information on those who have attrited (Behrman, Parker and Todd, 2009; Jensen, 2010; Duflo, Hanna and Ryan, 2012). The outcome of interest is constructed using observed information on individuals surveyed and information reported by others for the sample of attritors. In this case the main concern is the reliability of the proxy reports and whether reliability is correlated with migration patterns in space and time. Reliability can be partly verified if double information exists on some migrants. Rosenzweig (2003), for instance, uses double information for those who migrated inside the village (self-reported and reported by other members) to validate proxy information on schooling outcomes for attritors in the Bangladesh Nutrition Survey (1981 to 2000). This test relies on the assumption that reliability using information reported by other household members living in the same

[43] Hull (2015) uses randomized intensive follow-up to overcome differential attrition in estimating Local Average Treatment Effects. He proposes using randomized assignment to intensive follow-up as a pre-randomization stratification variable. Since intensive surveying is random, this stratification is likely uncorrelated with the distribution of complier treatment effects.

village, is relevant for proxy reports on far away migrants, for whom outcomes may be harder to observe by prior household members.

Dealing with attrition *ex-post*

Even after intensive tracking some attrition will almost always remain, which can be non-random. When attrition causes samples to become unbalanced, adjusting for covariate differences may remove biases, even if one generally may want to limit controls in ITT estimates of a randomized assigned intervention (Athey and Imbens, 2017). The econometrics literature further proposes several alternative methods to acquire consistent estimates in the presence of non-random missing data (Heckman, 1979; Rubin, 1987; Robins, Rotnitzky and Zhao, 1995; Wooldridge, 2002*a*), depending on the nature of the selection process. Fitzgerald, Gottschalk and Moffitt (1998) distinguish between identifiability under selection on observables and on unobservables. If attrition is driven by selection on observables, unbiased estimates can be obtained using weighted least square regression (Fitzgerald, Gottschalk and Moffitt, 1998; Wooldridge, 2002*a*).

In case of non-random selection driven by unobservables, a Heckman sample selection correction model can be used if there is a credible "exclusion restriction". But finding variables that are completely exogenous from the outcome of interest but highly correlated to the probability of being found can be challenging. A set of credible exogenous variables are sometimes formed by the characteristics of the survey and tracking design (Zabel, 1998; Hill and Willis, 2001). Maluccio (2004) uses information reflecting the quality of the fieldwork during the first round of KIDS to correct for attrition bias on follow-up rounds. Thomas et al. (2012) use information from a Survey of surveyors conducted during the second wave of the IFLS to predict survey status in later waves of data. Interviewer characteristics can be used as instruments in a selection model, but only if they are not correlated with respondent characteristics.[44]

[44]This implies that ideally interviewers should be randomly assigned. We found only two studies in which a Heckman Selection Model was used to correct for attrition using information on a randomized survey design (Dinkelman and Martínez A, 2014; Fitzsimons et al., 2016).

As both IPW and Heckman's correction selection model are based on strong assumptions, it has become relatively common in the impact evaluation literature to use instead non-parametric techniques to bound estimates. Depending on the outcome of interest, different types of bounds can be estimated. For bounded outcomes, Horowitz and Manski (2000) proposed to construct bounds by assuming that those who are missing represents the "worst cases" and missing information is imputed using minimal and maximal possible values of the outcome variables. Therefore, the outcome variable has to be bounded but no assumption on the selection mechanisms are needed. While bounds can provide useful benchmarks for binary outcomes, for outcomes with wide support, the bounds can be very wide and non-informative. To relax the extreme assumption on the distribution of treatment effects among attritors, Kling, Liebman and Katz (2007) construct bounds using the mean and standard deviation of the observed treatment and control distribution. Hence, they propose an alternative assumption about positive (negative) attrition bias based on treated attritors being below (above) the observed treatment mean by a half standard deviation and control attritors being above (below) the observed experimental control mean by half a standard deviation. This specification leads to tighter intervals by assuming that attritors in each experimental group behave somewhat similar to observed individual of that group.

Finally, Lee (2009) proposes to bound the treatment estimate for those who are always observed whenever attrition is not balanced between treatment groups. Instead of constructing a worst-case scenario, bounds are estimated by trimming a share of the sample, either from above or from below. To apply this type of bounds, two assumptions need to be satisfied. First, the treatment has to be randomly assigned and second, assignment to treatment can only affect attrition in one direction (monotonicity assumption). To obtain tighter bounds, lower and upper bounds can be estimated using a small number of covariates and trimming the sample by cell. Lee bounds are relatively often used to correct for attrition in RCTs.

At the intersection between Lee bounds and Heckman sample selectivity correction models, Behaghel et al. (2015) use the number of attempts to obtain responses to a survey from each respondent as an instrument of sample selection. They present a semi-parametric version of Heckman's latent selection model, in which respondents are ranked by their reluctance to respond. This approach truncates the sample of respondents in the treatment arm with higher response rate using as benchmark the number of attempts needed to acquire the same share of respondents in both groups, to restore balance after sample selection and get a local estimate of treatment effects. As for Lee bounds, this approach requires the monotonicity condition on response behavior, but in this case the monotonicity condition should hold jointly on the impact of assignment to treatment and on the impact of survey effort.

G.2 Attrition rates in RCTs

In this section we reviewed the literature to document how development economics papers handle attrition.

To assess a representative sample of high-quality papers the review was limited to articles published in top economic journals: the American Economic Review, Journal of Development Economics, The Quarterly Journal of Economics, the American Economic Journal: Applied Economics, the Journal of Political Economy, The Economic Journal, Econometrica and The Review of Economics and Statistics. The search covered articles published from 2009 to the first quarter of 2019.[45] In the first step we identify those articles in which the identification strategy exploits a randomized implementation of an intervention in a developing country. We limited the review to developing countries. In the second step, we keep only articles satisfying the following conditions:

- The intervention targeted individuals or households. We also keep interventions targeting schools if the final unit of analysis was the student.
- Data was collected through a household or an individual survey. We drop papers relying only on administrative data, diaries or logbooks.

[45] Conclusions from the literature review are broadly similar when focusing on a somewhat earlier period, 2009 to 2015 (see Molina-Millán and Macours (2017)). Similar conclusions are also drawn in a recent independent review by Hirshleifer, Ortiz Becerra and Ghanem (2019).

- The analysis uses at least two rounds of panel data.

The final sample includes 144 articles (see below).

Table G1 presents descriptive statistics on attrition rates. We divide studies into three categories depending on the unit of analysis: children under 18 years old (39 studies), households (51 studies) and adults (62 studies). For four papers we could not find information on the attrition rate. Eight studies report attrition rates for different sub-samples of respondents. Table G1 shows large variation in attrition rates. As expected, studies targeting households have on average lower attrition rates than studies targeting individuals, and attrition rates are higher for individual adults than for children. In general, articles in this last category include mainly studies targeting young adults who tend to be more mobile. Table G2 shows descriptive statistics of accumulated attrition rates and shows a similar pattern. Notably, about half of the studies estimating outcomes for individual adults have attrition rates of 13.8 percent or higher.

Most studies using panel data to evaluate a RCT report attrition incidence by treatment arms: still almost 18 percent of the studies do not include this basic information. Leaving those exceptions aside, the general practice is to report whether attrition rates are balanced between treatment arms (including between treatment and control). Seventy four percent of the studies that test for differences, could not reject the null hypothesis that attrition rates are balanced between treatment arms. This implies that non-ignorable share of studies had to deal with significant differences on attrition rates among treatments arms. In most cases the authors resort to restricting the analysis to a round of data, a subsample, a region or a particular set of treatments for which attrition was balanced.

There notably is a large heterogeneity in how studies account for attrition after this first step, and overall, often limited consideration of potential attrition bias - see Table G3. In 23 percent of the studies that reported attrition rates by treatment arms the analysis of attrition is limited to this first step. The rest of the studies also analyzes how attritors differ from those who stay. Hence, 28 percent of the studies report whether baseline characteristics for the subsample of respondents are balanced after attrition, and 45 percent consider whether selection into attrition is driven differently by baseline characteristics among treatment groups (in some cases the authors include treatment interactions). In 25 percent of the studies the authors report that attrition is not random. However, in many cases the authors only look at the outcome of interest at baseline or at a small list of baseline characteristics like gender or age.

Among the 36 studies that report non-random attrition 11 do not apply any method to correct for non-random attrition - see Table G5. The other mainly apply IPW (5), non-parametric bounds (15) or both (4). In another six cases the authors report IPW even if they have not detected non-random attrition. Many authors acknowledge IPW only permits to correct for selection on observables, while selection on unobservables may still bias the results. It is becoming more common in the literature to report bounds for the range of treatment effects. Almost 23 percent of the articles reviewed show estimates on the upper and lower bound following the methodologies in Horowitz and Manski (2000); Lee (2002); Kling, Liebman and Katz (2007). While in fifty percent of the studies reporting bounds we only find bounds as in Lee (2002, 2009), it is common to present bounds constructed making different assumptions on the distribution of treatment effects. Thirty percent of the studies report worst-case scenario bounds (Horowitz and Manski, 2000) while 21 percent presents less extreme bounds as in Kling, Liebman and Katz (2007) which depending on the assumptions made could end being very narrow. Finally, 8 percent report bounds following the three methodologies.

Comparing Tables G4 and G6 with Tables G3 and G5, further shows that limiting the analysis to studies published in the last 5 years leads to qualitatively similar findings. This confirms that the variation in the way studies report and account for attrition is also observed in recent work.

Table G1: Annual attrition rates in the sample studies

	Children under 18 years old	Household	Adults above 18 years old
Average	9.20	7.50	16.07
Standard Deviation	10.63	7.67	18.51
Minimum	0.6	0	0
Median	6.28	5.11	9.17
Maximum	60	30	75
Number of studies	38	46	57

Note: We exclude 7 studies for which follow-up data was collected less than 3 months after baseline. In 8 studies the authors report attrition rates for different subsamples of respondents, in such cases we treat each target group as a separately study and we include in the statistics both attrition rates.

Table G2: Accumulated attrition rates in the sample studies

	Children under 18 years old	Household	Adults above 18 years old
Average	14.17	9.18	16.02
Standard Deviation	10.80	7.48	13.61
Minimum	1.5	0	0
Median	11.3	6.7	13.8
Maximum	40	25	67.42
Number of studies	38	49	61

Note: In 8 studies the authors report attrition rates for different subsamples of respondents, in such cases we treat each target group as a separately study and we include in the statistics both attrition rates.

Table G3: Reporting and dealing with attrition in studies published in 2009-2019 (142 studies)

				Analyze		
	Report Attrition by Treatment Arms (mean comparison)	Find Balanced Attrition Rates (t-test)	Do not Proceed any Further	Correlates of Attrition	Balanced Randomization After Atrittion	Heterogeneous Treatment Effects
Number of studies	117	87	20	62	40	64
% of total studies	82.39	61.23	14.08	43.66	28.17	45.07
% of studies col. 1		74.36	22.99			

Note: The total number of studies does not include 6 studies in which the reported attrition rate is 0. Studies in which authors analyze two different subsamples of respondents are counted double.

Table G4: Reporting and dealing with attrition in studies published after 2015 (80 studies)

	Report Attrition by Treatment Arms (mean comparison)	Find Balanced Attrition Rates (t-test)	Do not Proceed any Further	Analyze		
				Correlates of Attrition	Balanced Randomization After Atrittion	Heterogeneous Treatment Effects
Number of studies	64	47	8	39	28	44
% of total studies	80	68.75	10	48.75	35	55
% of studies col. 1		73.44	17.02			

Note: The total number of studies does not include 6 studies in which the reported attrition rate is 0. Studies in which authors analyze two different subsamples of respondents are counted double.

Table G5: Methods to correct for attrition in the studies published in 2009-2019 (142 studies)

	IPW	Non-parametric Bounds	IPW+ Bounds	Proxy Information	Heckman Correction model	Proxy+Heckman Correction model	None
Number of studies	13	29	5	2	1	1	102
Selection on observables:							
Random Attrition	5	7	1	1	0	0	31
Non-random Attrition	5	15	4	0	1	0	11
No reported	3	7	0	1	0	1	60
% of total studies	9.15	20.42	3.52	1.41	0.7	0.7	71.83

Note: The total number of studies does not include 6 studies in which the reported attrition rate is 0. Studies in which authors analyze two different subsamples of respondents are counted double.

Table G6: Methods to correct for attrition in studies published after 2015 (80 studies)

	IPW	Non-parametric Bounds	IPW+ Bounds	Proxy Information	Heckman Correction model	Proxy+Heckman Correction model	None
Number of studies	7	21	3	1	1	0	58
Selection on observables:							
Random Attrition	1	5	1	1	0	0	21
Non-random Attrition	4	14	2	0	1	0	4
No reported	2	2	0	0	0	0	33
% of total studies	8.75	26.25	3.75	1.25	1.25	0	72.5

Note: The total number of studies does not include 6 studies in which the reported attrition rate is 0. Studies in which authors analyze two different subsamples of respondents are counted double.

List of papers in the review

Afzal, Uzma, Giovanna d'Adda, Marcel Fafchamps, Simon Quinn, and Farah Said. 2017. "Two Sides of the Same Rupee? Comparing Demand for Microcredit and Microsaving in a Framed field experiment in rural Pakistan." *The Economic Journal*, 128(614): 2161–2190.

Aggarwal, Shilpa, Eilin Francis, and Jonathan Robinson. 2018. "Grain Today, Gain Tomorrow: Evidence from a Storage Experiment with Savings Clubs in Kenya." *Journal of Development Economics*, 134: 1 – 15.

Aguila, Emma, Arie Kapteyn, and Francisco Perez-Arce. 2017. "Consumption Smoothing and Frequency of Benefit Payments of Cash Transfer Programs." *American Economic Review*, 107(5): 430–35.

Aker, Jenny C, Christopher Ksoll, and Travis J Lybbert. 2012. "Can Mobile Phones Improve Learning? Evidence from a Field Experiment in Niger." *American Economic Journal: Applied Economics*, 4(4): 94–120.

Alan, Sule, and Seda Ertac. 2018. "Fostering Patience in the Classroom: Results from Randomized Educational Intervention." *Journal of Political Economy*, 126(5): 1865–1911.

Alan, Sule, Teodora Boneva, and Seda Ertac. 2019. "Ever Failed, Try Again, Succeed Better: Results from a Randomized Educational Intervention on Grit." *The Quarterly Journal of Economics*, 134(3): 1121–1162.

Andrabi, Tahir, Jishnu Das, and Asim Ijaz Khwaja. 2017. "Report Cards: The Impact of Providing School and Child Test Scores on Educational Markets." *American Economic Review*, 107(6): 1535–63.

Angelucci, Manuela. 2015. "Migration and Financial Constraints: Evidence from Mexico." *The Review of Economics and Statistics*, 97(1): 224–228.

Angelucci, Manuela, Dean Karlan, and Jonathan Zinman. 2015. "Microcredit Impacts: Evidence from a Randomized Microcredit Program Placement Experiment by Compartamos Banco." *American Economic Journal: Applied Economics*, 7(1): 151–182.

Araujo, M. Caridad, Pedro Carneiro, Yyannú Cruz-Aguayo, and Norbert Schady. 2016. "Teacher Quality and Learning Outcomes in Kindergarten." *The Quarterly Journal of Economics*, 131(3): 1415–1453.

Ashraf, Nava, Diego Aycinena, Claudia Martínez, and Dean Yang. 2015. "Savings in Transnational Households: a Field Experiment among Migrants from El Salvador." *Review of Economics and Statistics*, 97(2): 332–351.

Ashraf, Nava, Erica Field, and Jean Lee. 2014. "Household Bargaining and Excess Fertility: an Experimental Study in Zambia." *The American Economic Review*, 104(7): 2210–2237.

Ashraf, Nava, James Berry, and Jesse M Shapiro. 2010. "Can Higher Prices Stimulate Product Use? Evidence from a Field Experiment in Zambia." *The American Economic Review*, 100(5): 2383–2413.

Attanasio, Orazio, Britta Augsburg, Ralph De Haas, Emla Fitzsimons, and Heike Harmgart. 2015. "The Impacts of Microfinance: Evidence from Joint-Liability Lending in Mongolia." *American Economic Journal: Applied Economics*, 7(1): 90–122.

Augsburg, Britta, Ralph De Haas, Heike Harmgart, and Costas Meghir. 2015. "The Impacts of Microcredit: Evidence from Bosnia and Herzegovina." *American Economic Journal: Applied Economics*, 7(1): 183–203.

Avitabile, Ciro, and Rafael de Hoyos. 2018. "The Heterogeneous Effect of Information on Student Performance: Evidence from a Randomized Control Trial in Mexico." *Journal of Development Economics*, 135: 318 – 348.

Baird, Sarah, Craig McIntosh, and Berk Özler. 2011. "Cash or Condition? Evidence from a Cash Transfer Experiment." *The Quarterly Journal of Economics*, 126(4): 1709–1753.

Baird, Sarah, Joan Hamory Hicks, Michael Kremer, and Edward Miguel. 2016. "Worms at Work: Long-run Impacts of a Child Health Investment." *The Quarterly Journal of Economics*, 131(4): 1637–1680.

Bandiera, Oriana, Robin Burgess, Narayan Das, Selim Gulesci, Imran Rasul, and Munshi Sulaiman. 2017. "Labor Markets and Poverty in Village Economies." *The Quarterly Journal of Economics*, 132(2): 811–870.

Banerjee, Abhijit, Esther Duflo, and Richard Hornbeck. 2014. "Bundling Health Insurance and Microfinance in India: There Cannot be Adverse Selection if There is No Demand." *The American Economic Review*, 104(5): 291–297.

Banerjee, Abhijit, Esther Duflo, Rachel Glennerster, and Cynthia Kinnan. 2015. "The Miracle of Microfinance? Evidence from a Randomized Evaluation." *American Economic Journal: Applied Economics*, 7(1): 22–53.

Banerjee, Abhijit, Sharon Barnhardt, and Esther Duflo. 2018. "Can Iron-fortified Salt Control Anemia? Evidence from Two Experiments in Rural Bihar." *Journal of Development Economics*, 133: 127–146.

Banerji, Rukmini, James Berry, and Marc Shotland. 2017. "The Impact of Maternal Literacy and Participation Programs: Evidence from a Randomized Evaluation in India." *American Economic Journal: Applied Economics*, 9(4): 303–37.

Barham, Tania, Karen Macours, and John A Maluccio. 2013. "Boys' Cognitive Skill Formation and Physical Growth: Long-Term Experimental Evidence on Critical Ages for Early Childhood Interventions." *The American Economic Review*, 103(3): 467–471.

Barnhardt, Sharon, Erica Field, and Rohini Pande. 2017. "Moving to Opportunity or Isolation? Network Effects of a Randomized Housing Lottery in Urban India." *American Economic Journal: Applied Economics*, 9(1): 1–32.

Basu, Karna, and Maisy Wong. 2015. "Evaluating Seasonal Food Storage and Credit Programs in East Indonesia." *Journal of Development Economics*, 115: 200 – 216.

Bauchet, Jonathan, Jonathan Morduch, and Shamika Ravi. 2015. "Failure vs. Displacement: Why an Innovative Anti-poverty Program Showed no Net Impact in South India." *Journal of Development Economics*, 116: 1 – 16.

Beaman, Lori, Dean Karlan, Bram Thuysbaert, and Christopher Udry. 2013. "Profitability of Fertilizer: Experimental Evidence from Female Rice Farmers in Mali." *The American Economic Review*, 103(3): 381–386.

Beam, Emily A. 2016. "Do Job Fairs Matter? Experimental Evidence on the Impact of Job-fair Attendance." *Journal of Development Economics*, 120: 32 – 40.

Beegle, Kathleen, Emanuela Galasso, and Jessica Goldberg. 2017. "Direct and Indirect Effects of Malawi's Public Works Program on Food Security." *Journal of Development Economics*, 128: 1 – 23.

Behrman, Jere R, Susan W Parker, Petra E Todd, and Kenneth I Wolpin. 2015. "Aligning Learning Incentives of Students and Teachers: Results from a Social Experiment in Mexican High Schools." *Journal of Political Economy*, 123(2): 325–364.

Beuermann, Diether W., Julian Cristia, Santiago Cueto, Ofer Malamud, and Yyannu Cruz-Aguayo. 2015. "One Laptop per Child at Home: Short-Term Impacts from a Randomized Experiment in Peru." *American Economic Journal: Applied Economics*, 7(2): 53–80.

Björkman Nyqvist, Martina, and Seema Jayachandran. 2017. "Mothers Care More, but Fathers Decide: Educating Parents about Child Health in Uganda." *American Economic Review*, 107(5): 496–500.

Björkman Nyqvist, Martina, Lucia Corno, Damien de Walque, and Jakob Svensson. 2018. "Incentivizing Safer Sexual Behavior: Evidence from a Lottery Experiment on HIV Prevention." *American Economic Journal: Applied Economics*, 10(3): 287–314.

Blattman, Christopher, and Stefan Dercon. 2018. "The Impacts of Industrial and Entrepreneurial Work on Income and Health: Experimental Evidence from Ethiopia." *American Economic Journal: Applied Economics*, 10(3): 1–38.

Blattman, Christopher, Eric P. Green, Julian Jamison, M. Christian Lehmann, and Jeannie Annan. 2016. "The Returns to Microenterprise Support among the Ultrapoor: A Field Experiment in Postwar Uganda." *American Economic Journal: Applied Economics*, 8(2): 35–64.

Blattman, Christopher, Julian C. Jamison, and Margaret Sheridan. 2017. "Reducing Crime and Violence: Experimental Evidence from Cognitive Behavioral Therapy in Liberia." *American Economic Review*, 107(4): 1165–1206.

Blattman, Christopher, Mathilde Emeriau, and Nathan Fiala. 2018. "Do Anti-Poverty Programs Sway Voters? Experimental Evidence from Uganda." *The Review of Economics and Statistics*, 100(5): 891–905.

Blattman, Christopher, Nathan Fiala, and Sebastian Martinez. 2014. "Generating Skilled Self-Employment in Developing Countries: Experimental Evidence from Uganda." *The Quarterly Journal of Economics*, 129(2): 697–752.

Blumenstock, Joshua, Michael Callen, and Tarek Ghani. 2018. "Why Do Defaults Affect Behavior? Experimental Evidence from Afghanistan." *American Economic Review*, 108(10): 2868–2901.

Breza, Emily, and Arun G. Chandrasekhar. 2019. "Social Networks, Reputation, and Commitment: Evidence From a Savings Monitors Experiment." *Econometrica*, 87(1): 175–216.

Brooks, Wyatt, Kevin Donovan, and Terence R. Johnson. 2018. "Mentors or Teachers? Microenterprise Training in Kenya." *American Economic Journal: Applied Economics*, 10(4): 196–221.

Bruhn, Miriam, Gabriel Lara Ibarra, and David McKenzie. 2014. "The Minimal Impact of a Large-Scale Financial Education Program in Mexico City." *Journal of Development Economics*, 108: 184–189.

Bruhn, Miriam, Luciana de Souza Leão, Arianna Legovini, Rogelio Marchetti, and Bilal Zia. 2016. "The Impact of High School Financial Education: Evidence from a Large-Scale Evaluation in Brazil." *American Economic Journal: Applied Economics*, 8(4): 256–95.

Brune, Lasse, Xavier Giné, Jessica Goldberg, and Dean Yang. 2017. "Savings Defaults and Payment Delays for Cash Transfers: Field Experimental Evidence from Malawi." *Journal of Development Economics*, 129(C): 1–13.

Bryan, Gharad, Shyamal Chowdhury, and Ahmed Mushfiq Mobarak. 2014. "Underinvestment in a Profitable Technology: The Case of Seasonal Migration in Bangladesh." *Econometrica*, 82(5): 1671–1748.

Burchardi, Konrad B, Selim Gulesci, Benedetta Lerva, and Munshi Sulaiman. 2018. "Moral Hazard: Experimental Evidence from Tenancy Contracts." *The Quarterly Journal of Economics*, 134(1): 281–347.

Burde, Dana, and Leigh L Linden. 2013. "Bringing Education to Afghan Girls: A Randomized Controlled Trial of Village-Based Schools." *American Economic Journal: Applied Economics*, 5(3): 27–40.

Burlando, Alfredo, and Andrea Canidio. 2017. "Does Group Inclusion Hurt Financial Inclusion? Evidence from Ultra-poor Members of Ugandan Savings Groups." *Journal of Development Economics*, 128(C): 24–48.

Busso, Matias, and Sebastian Galiani. 2019. "The Causal Effect of Competition on Prices and Quality: Evidence from a Field Experiment." *American Economic Journal: Applied Economics*, 11(1): 33–56.

Carvalho, Leandro S., Silvia Prina, and Justin Sydnor. 2016. "The Effect of Saving on Risk Attitudes and Intertemporal Choices." *Journal of Development Economics*, 120: 41 – 52.

Chinkhumba, Jobiba, Susan Godlonton, and Rebecca L Thornton. 2016. "The Demand for Medical Male Circumcision." *American Economic Journal: Applied Economics*, 6(2): 152–177.

Chong, Alberto, Isabelle Cohen, Erica Field, Eduardo Nakasone, and Maximo Torero. 2016. "Iron Deficiency and Schooling Attainment in Peru." *American Economic Journal: Applied Economics*, 8(4): 222–55.

Cohen, Jessica, and Indrani Saran. 2018. "The Impact of Packaging and Messaging on Adherence to Malaria Treatment: Evidence from a Randomized Controlled Trial in Uganda." *Journal of Development Economics*, 134: 68 – 95.

Cohen, Jessica, and Pascaline Dupas. 2010. "Free Distribution or Cost-Sharing? Evidence from a Randomized Malaria Prevention Experiment." *The Quarterly Journal of Economics*, 125(1): 1–45.

Cohen, Jessica, Katherine Lofgren, and Margaret McConnell. 2017. "Precommitment, Cash Transfers, and Timely Arrival for Birth: Evidence from a Randomized Controlled Trial in Nairobi Kenya." *American Economic Review*, 107(5): 501–05.

Cohen, Jessica, Pascaline Dupas, and Simone Schaner. 2015. "Price Subsidies, Diagnostic Tests, and Targeting of Malaria Treatment: Evidence from a Randomized Controlled Trial." *The American Economic Review*, 105(2): 609–645.

Cole, Shawn, Daniel Stein, and Jeremy Tobacman. 2014. "Dynamics of Demand for Index Insurance: Evidence from a Long-Run Field Experiment." *The American Economic Review*, 104(5): 284–290.

Collier, Paul, and Pedro C Vicente. 2014. "Votes and Violence: Evidence from a Field Experiment in Nigeria." *The Economic Journal*, 124(574): F327–F355.

Crépon, Bruno, Florencia Devoto, Esther Duflo, and William Parienté. 2015. "Estimating the Impact of Microcredit on Those Who Take It Up: Evidence from a Randomized Experiment in Morocco." *American Economic Journal: Applied Economics*, 7(1): 123–150.

Cristia, Julian, Pablo Ibarrarán, Santiago Cueto, Ana Santiago, and Eugenio Severín. 2017. "Technology and Child Development: Evidence from the One Laptop per Child Program." *American Economic Journal: Applied Economics*, 9(3): 295–320.

Cunha, Jesse M. 2014. "Testing Paternalism: Cash Versus In-Kind Transfers." *American Economic Journal: Applied Economics*, 6(2): 195–230.

Das, Jishnu, Stefan Dercon, James Habyarimana, Pramila Krishnan, Karthik Muralidharan, and Venkatesh Sundararaman. 2013. "School Inputs, Household Substitution, and Test Scores." *American Economic Journal: Applied Economics*, 5(2): 29–57.

de Ree, Joppe, Karthik Muralidharan, Menno Pradhan, and Halsey Rogers. 2018. "Double for Nothing? Experimental Evidence on an Unconditional Teacher Salary Increase in Indonesia." *The Quarterly Journal of Economics*, 133(2): 993–1039.

Dhaliwal, Iqbal, and Rema Hanna. 2017. "The Devil is in the Details: The Successes and Limitations of Bureaucratic Reform in India." *Journal of Development Economics*, 124: 1 – 21.

Dinkelman, Taryn, and Claudia Martínez A. 2014. "Investing in Schooling in Chile: The Role of Information About Financial Aid for Higher Education." *Review of Economics and Statistics*, 96(2): 244–257.

Doi, Yoko, David McKenzie, and Bilal Zia. 2014. "Who You Train Matters: Identifying Combined Effects of Financial Education on Migrant Households." *Journal of Development Economics*, 109: 39–55.

Drexler, Alejandro, Greg Fischer, and Antoinette Schoar. 2014. "Keeping it Simple: Financial Literacy and Rules of Thumb." *American Economic Journal: Applied Economics*, 6(2): 1–31.

Duflo, Esther, Michael Kremer, and Jonathan Robinson. 2011. "Nudging Farmers to Use Fertilizer: Theory and Experimental Evidence from Kenya." *The American Economic Review*, 101(6): 2350–2390.

Duflo, Esther, Pascaline Dupas, and Michael Kremer. 2011. "Peer effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya." *The American Economic Review*, 101(5): 1739–1774.

Duflo, Esther, Pascaline Dupas, and Michael Kremer. 2015. "Education, HIV, and Early Fertility: Experimental Evidence from Kenya." *American Economic Review*, 105(9): 2757–97.

Duflo, Esther, Rema Hanna, and Stephen P Ryan. 2012. "Incentives Work: Getting Teachers to Come to School." *The American Economic Review*, 102(4): 1241–1278.

Dupas, Pascaline. 2014. "Short-Run Subsidies and Long-Run Adoption of New Health Products: Evidence from a Field Experiment." *Econometrica*, 82(1): 197–228.

Dupas, Pascaline, and Jonathan Robinson. 2013. "Why Don't the Poor Save More? Evidence from Health Savings Experiments." *American Economic Review*, 103(4): 1138–71.

Dupas, Pascaline, Anthony Keats, and Jonathan Robinson. 2017. "The Effect of Savings Accounts on Interpersonal Financial Relationships: Evidence from a Field Experiment in Rural Kenya." *The Economic Journal*, 129(617): 273–310.

Dupas, Pascaline, Dean Karlan, Jonathan Robinson, and Diego Ubfal. 2018. "Banking the Unbanked? Evidence from Three Countries." *American Economic Journal: Applied Economics*, 10(2): 257–97.

Edmonds, Eric V, and Maheshwor Shrestha. 2014. "You Get What You Pay For: Schooling Incentives and Child Labor." *Journal of Development Economics*, 111: 196–211.

Field, Erica, Rohini Pande, John Papp, and Natalia Rigol. 2013. "Does the Classic Microfinance Model Discourage Entrepreneurship Among the Poor? Experimental Evidence from India." *The American Economic Review*, 103(6): 2196–2226.

Field, Erica, Seema Jayachandran, and Rohini Pande. 2010. "Do Traditional Institutions Constrain Female Entrepreneurship? A Field Experiment on Business Training in India." *The American Economic Review*, 100(2): 125–129.

Fitzsimons, Emla, Bansi Malde, Alice Mesnard, and Marcos Vera-Hernandez. 2016. "Nutrition, Information and Household Behavior: Experimental Evidence from Malawi." *Journal of Development Economics*, 122: 113 – 126.

Flory, Jeffrey A. 2018. "Formal Finance and Informal Safety Nets of the Poor: Evidence from a Savings Field Experiment." *Journal of Development Economics*, 135: 517 – 533.

Franklin, Simon. 2017. "Location, Search Costs and Youth Unemployment: Experimental Evidence from Transport Subsidies." *The Economic Journal*, 128(614): 2353–2379.

Gertler, Paul J, Sebastian W Martinez, and Marta Rubio-Codina. 2012. "Investing Cash Transfers to Raise Long-Term Living Standards." *American Economic Journal: Applied Economics*, 4(1): 164–192.

Giné, Xavier, and Ghazala Mansuri. 2018. "Together We Will: Experimental Evidence on Female Voting Behavior in Pakistan." *American Economic Journal: Applied Economics*, 10(1): 207–35.

Giné, Xavier, Jessica Goldberg, and Dean Yang. 2012. "Credit Market Consequences of Improved Personal Identification: Field Experimental Evidence from Malawi." *The American Economic Review*, 102(6): 2923–2954.

Giné, Xavier, Jessica Goldberg, Dan Silverman, and Dean Yang. 2018. "Revising Commitments: Field Evidence on the Adjustment of Prior Choices." *Economic Journal*, 128(608): 159–188.

Glewwe, Paul, Albert Park, and Meng Zhao. 2016. "A Better Vision for Development: Eyeglasses and Academic Performance in Rural Primary Schools in China." *Journal of Development Economics*, 122(C): 170–182.

Glewwe, Paul, Nauman Ilias, and Michael Kremer. 2010. "Teacher Incentives." *American Economic Journal: Applied Economics*, 2(3): 205–227.

Godlonton, Susan, Alister Munthali, and Rebecca Thornton. 2016. "Responding to Risk: Circumcision, Information, and HIV Prevention." *The Review of Economics and Statistics*, 98(2): 333–349.

Godlonton, Susan, and Rebecca L Thornton. 2013. "Learning from Others' HIV Testing: Updating Beliefs and Responding to Risk." *The American Economic Review*, 103(3): 439–444.

Gong, Erick. 2015. "HIV Testing and Risky Sexual Behaviour." *The Economic Journal*, 125(582): 32–60.

Groh, Matthew, and David McKenzie. 2016. "Macroinsurance for Microenterprises: A Randomized Experiment in Post-revolution Egypt." *Journal of Development Economics*, 118: 13 – 25.

Groh, Matthew, Nandini Krishnan, David McKenzie, and Tara Vishwanath. 2016. "Do Wage Subsidies Provide a Stepping-Stone to Employment for Recent College Graduates? Evidence from a Randomized Experiment in Jordan." *The Review of Economics and Statistics*, 98(3): 488–502.

Hanna, Rema, and Paulina Oliva. 2015. "Moving Up the Energy Ladder: The Effect of an Increase in Economic Well-Being on the Fuel Consumption Choices of the Poor in India." *American Economic Review*, 105(5): 242–46.

Haushofer, Johannes, and Jeremy Shapiro. 2016. "The Short-Term Impact of Unconditional Cash Transfers to the Poor: Experimental Evidence from Kenya." *The Quarterly Journal of Economics*, 131: qjw025.

Hicken, Allen, Stephen Leider, Nico Ravanilla, and Dean Yang. 2015. "Measuring Vote-Selling: Field Evidence from the Philippines." *American Economic Review*, 105(5): 352–56.

Hicken, Allen, Stephen Leider, Nico Ravanilla, and Dean Yang. 2018. "Temptation in Vote-selling: Evidence from a Field Experiment in the Philippines." *Journal of Development Economics*, 131(C): 1–14.

Hidrobo, Melissa, Amber Peterman, and Lori Heise. 2016. "The Effect of Cash, Vouchers, and Food Transfers on Intimate Partner Violence: Evidence from a Randomized Experiment in Northern Ecuador." *American Economic Journal: Applied Economics*, 8(3): 284–303.

Hidrobo, Melissa, John Hoddinott, Amber Peterman, Amy Margolies, and Vanessa Moreira. 2014. "Cash, Food, or Vouchers? Evidence from a Randomized Experiment in Northern Ecuador." *Journal of Development Economics*, 107: 144–156.

Hirshleifer, Sarojini, David McKenzie, Rita Almeida, and Cristobal Ridao-Cano. 2016. "The Impact of Vocational Training for the Unemployed: Experimental Evidence from Turkey." *Economic Journal*, 126(597): 2115–2146.

Jensen, Robert. 2010. "The (Perceived) Returns to Education and the Demand for Schooling." *The Quarterly Journal of Economics*, 125(2): 515–548.

Jensen, Robert. 2012. "Do Labor Market Opportunities Affect Young Women's Work and Family Decisions? Experimental Evidence from India." *The Quarterly Journal of Economics*, 127(2): 753–792.

Jensen, Robert T, and Nolan H Miller. 2011. "Do Consumer Price Subsidies Really Improve Nutrition?" *Review of Economics and Statistics*, 93(4): 1205–1223.

Jones, Maria, and Florence Kondylis. 2018. "Does Feedback Matter? Evidence from Agricultural Services." *Journal of Development Economics*, 131(C): 28–41.

Karlan, Dean, and Martin Valdivia. 2011. "Teaching Entrepreneurship: Impact of Business Training on Microfinance Clients and Institutions." *Review of Economics and Statistics*, 93(2): 510–527.

Kast, Felipe, Stephan Meier, and Dina Pomeranz. 2018. "Saving More in Groups: Field Experimental Evidence from Chile." *Journal of Development Economics*, 133: 275 – 294.

Kazianga, Harounan, Damien de Walque, and Harold Alderman. 2014. "School Feeding Programs, Intrahousehold Allocation and the Nutrition of Siblings: Evidence from a Randomized Trial in Rural Burkina Faso." *Journal of Development Economics*, 106: 15–34.

Kondylis, Florence, Valerie Mueller, and Jessica Zhu. 2017. "Seeing is Believing? Evidence from an Extension Network Experiment." *Journal of Development Economics*, 125: 1 – 20.

Kremer, Michael, Edward Miguel, and Rebecca Thornton. 2009. "Incentives to Learn." *The Review of Economics and Statistics*, 91(3): 437–456.

Ksoll, Christopher, Helene Bie Lilleør, Jonas Helth Lønborg, and Ole Dahl Rasmussen. 2016. "Impact of Village Savings and Loan Associations: Evidence from a Cluster Randomized Trial." *Journal of Development Economics*, 120: 70 – 85.

Laajaj, Rachid. 2017. "Endogenous Time Horizon and Behavioral Poverty Trap: Theory and Evidence from Mozambique." *Journal of Development Economics*, 127(C): 187–208.

Labonne, Julien, and Robert S Chase. 2011. "Do Community-Driven Development Projects Enhance Social Capital? Evidence from the Philippines." *Journal of Development Economics*, 96(2): 348–358.

Lafortune, Jeanne, Julio Riutort, and José Tessada. 2018. "Role Models or Individual Consulting: The Impact of Personalizing Micro-entrepreneurship Training." *American Economic Journal: Applied Economics*, 10(4): 222–45.

León, Gianmarco. 2017. "Turnout, Political Preferences and Information: Experimental Evidence from Peru." *Journal of Development Economics*, 127(C): 56–71.

Levine, David, Rachel Polimeni, and Ian Ramage. 2016. "Insuring Health or Insuring Wealth? An Experimental Evaluation of Health Insurance in Rural Cambodia." *Journal of Development Economics*, 119: 1 – 15.

Luoto, Jill, David Levine, Jeff Albert, and Stephen Luby. 2014. "Nudging to Use: Achieving Safe Water Behaviors in Kenya and Bangladesh." *Journal of Development Economics*, 110: 13–21.

Macours, Karen, and Renos Vakis. 2014. "Changing Households' Investment Behaviour Through Social Interactions With Local Leaders: Evidence from a Randomised Transfer Programme." *The Economic Journal*, 124(576): 607–633.

Macours, Karen, Norbert Schady, and Renos Vakis. 2012. "Cash Transfers, Behavioral Changes, and Cognitive Development in Early Childhood: Evidence from a Randomized Experiment." *American Economic Journal: Applied Economics*, 4(2): 247–273.

Magnan, Nicholas, David J. Spielman, Travis J. Lybbert, and Kajal Gulati. 2015. "Leveling with Friends: Social Networks and Indian Farmers' Demand for a Technology with Heterogeneous Benefits." *Journal of Development Economics*, 116: 223 – 251.

Maitra, Pushkar, Sandip Mitra, Dilip Mookherjee, Alberto Motta, and Sujata Visaria. 2017. "Financing Smallholder Agriculture: An Experiment with Agent-intermediated Microloans in India." *Journal of Development Economics*, 127(C): 306–337.

Martínez A., Claudia, and Marcela Perticará. 2017. "Childcare Effects on Maternal Employment: Evidence from Chile." *Journal of Development Economics*, 126: 127 – 137.

Martínez A., Claudia, Esteban Puentes, and Jaime Ruiz-Tagle. 2018. "The Effects of Micro-entrepreneurship Programs on Labor Market Performance: Experimental Evidence from Chile." *American Economic Journal: Applied Economics*, 10(2): 101–24.

McKenzie, David. 2017. "Identifying and Spurring High-Growth Entrepreneurship: Experimental Evidence from a Business Plan Competition." *American Economic Review*, 107(8): 2278–2307.

Meredith, Jennifer, Jonathan Robinson, Sarah Walker, and Bruce Wydick. 2013. "Keeping the Doctor Away: Experimental Evidence on Investment in Preventative Health Products." *Journal of Development Economics*, 105: 196–210.

Mitra, Sandip, Dilip Mookherjee, Maximo Torero, and Sujata Visaria. 2018. "Asymmetric Information and Middleman Margins: An Experiment with Indian Potato Farmers." *The Review of Economics and Statistics*, 100(1): 1–13.

Muralidharan, Karthik, and Venkatesh Sundararaman. 2011. "Teacher Performance Pay: Experimental Evidence from India." *The Journal of Political Economy*, 119(1): 39–77.

Muralidharan, Karthik, and Venkatesh Sundararaman. 2015. "The Aggregate Effect of School Choice: Evidence from a Two-Stage Experiment in India." *The Quarterly Journal of Economics*, 130.

Olken, Benjamin A, Junko Onishi, and Susan Wong. 2014. "Should Aid Reward Performance? Evidence from a Field Experiment on Health and Education in Indonesia." *American Economic Journal: Applied Economics*, 6(4): 1–34.

Omotilewa, Oluwatoba J., Jacob Ricker-Gilbert, John Herbert Ainembabazi, and Gerald E. Shively. 2018. "Does Improved Storage Technology Promote Modern Input Use and Food Security? Evidence from a Randomized Trial in Uganda." *Journal of Development Economics*, 135: 176 – 198.

Powell-Jackson, Timothy, Kara Hanson, Christopher JM Whitty, and Evelyn K Ansah. 2014. "Who Benefits from Free Healthcare? Evidence from a Randomized Experiment in Ghana." *Journal of Development Economics*, 107: 305–319.

Prina, Silvia. 2015. "Banking the Poor Via Savings Accounts: Evidence from a Field Experiment." *Journal of Development Economics*, 115: 16–31.

Roy, Shalini, Jinnat Ara, Narayan Das, and Agnes R. Quisumbing. 2015. "“Flypaper Effects” in Transfers Targeted to Women: Evidence from BRAC's “Targeting the Ultra Poor” Program in Bangladesh." *Journal of Development Economics*, 117: 1 – 19.

Schaner, Simone. 2018. "The Persistent Power of Behavioral Change: Long-Run Impacts of Temporary Savings Subsidies for the Poor." *American Economic Journal: Applied Economics*, 10(3): 67–100.

Seshan, Ganesh, and Dean Yang. 2014. "Motivating Migrants: A field Experiment on Financial Decision-Making in Transnational Households." *Journal of Development Economics*, 108: 119–127.

Somville, Vincent, and Lore Vandewalle. 2018. "Saving by Default: Evidence from a Field Experiment in Rural India." *American Economic Journal: Applied Economics*, 10(3): 39–66.

Steinert, Janina Isabel, Lucie Dale Cluver, Franziska Meinck, Jenny Doubt, and Sebastian Vollmer. 2018. "Household Economic Strengthening Through Financial and Psychosocial Programming: Evidence from a Field Experiment in South Africa." *Journal of Development Economics*, 134: 443 – 466.

Szabó, Andrea, and Gergely Ujhelyi. 2015. "Reducing Nonpayment for Public Utilities: Experimental Evidence from South Africa." *Journal of Development Economics*, 117: 20 – 31.

Tarozzi, Alessandro, Aprajit Mahajan, Brian Blackburn, Dan Kopf, Lakshmi Krishnan, and Joanne Yoong. 2014. "Micro-Loans, Insecticide-Treated Bednets, and Malaria: Evidence from a Randomized Controlled Trial in Orissa, India." *The American Economic Review*, 104(7): 1909–1941.

Thornton, Rebecca L. 2012. "HIV Testing, Subjective Beliefs and Economic Behavior." *Journal of Development Economics*, 99(2): 300–313.

Valdivia, Martín. 2015. "Business Training Plus for Female Entrepreneurship? Short and Medium-Term Experimental Evidence from Peru." *Journal of Development Economics*, 113: 33–51.

Vicente, Pedro C. 2014. "Is Vote Buying Effective? Evidence from a Field Experiment in West Africa." *The Economic Journal*, 124(574): F356–F387.

Wilson, Nicholas L, Wentao Xiong, and Christine L Mattson. 2014. "Is Sex Like Driving? HIV Prevention and Risk Compensation." *Journal of Development Economics*, 106: 78–91.

Yi, Hongmei, Yingquan Song, Chengfang Liu, Xiaoting Huang, Linxiu Zhang, Yunli Bai, Baoping Ren, Yaojiang Shi, Prashant Loyalka, James Chu, et al. 2015. "Giving Kids a Head Start: The Impact and Mechanisms of Early Commitment of Financial Aid on Poor Students in Rural China." *Journal of Development Economics*, 113: 1–15.

G.3 Attrition rates in longitudinal surveys.

Table G7: Descriptive statistics: annual attrition rates in longitudinal surveys

	Same location	Village	Inside country	To other countries	Not Available
Average	12.91	1.27	2.25	4.45	2.07
Standard Deviation	3.99	1.1	1.86	5.43	2.55
Minimum	7.78	0	0.33	0.4	0.16
Median	12.84	1.18	1.72	2.71	0.62
Maximum	17.67	2.85	6.45	12	5.4
Number of surveys	6	6	9	4	6

Table G8: Descriptive statistics: accumulated attrition rates in longitudinal surveys

	Same location	Village	Inside country	To other countries	Not Available
Average	41.35	15.24	21.46	11.36	7.37
Standard Deviation	16.75	13.26	12.42	9.48	10.1
Minimum	28	0	2.6	1.45	0.64
Median	34.4	13.63	24	10	2.48
Maximum	70	37	38	24	26.5
Number of surveys	6	6	9	4	6

References Longitudinal Surveys

Alderman, Harold, Jere Behrman, Hans-Peter Kohler, John A Maluccio, and Susan Watkins. 2001. "Attrition in Longitudinal Household Survey Data: Some Tests from Three Developing Countries." *Demographic Research*, 5: 79–124.

Anglewicz, Philip, Jimi Adams, Francis Obare-Onyango, Hans-Peter Kohler, and Susan Watkins. 2009. "The Malawi Diffusion and Ideational Change Project 2004–06: Data Collection, Data Quality, and Analysis of Attrition." *Demographic Research*, 20(21): 503.

Beegle, Kathleen, Joachim De Weerdt, and Stefan Dercon. 2009. "The Intergenerational Impact of the African Orphans Crisis: a Cohort Study from an HIV/AIDS Affected Area." *International Journal of Epidemiology*, 38(2): 561–568.

Beegle, Kathleen, Joachim De Weerdt, and Stefan Dercon. 2011. "Migration and Economic Mobility in Tanzania: Evidence from a Tracking Survey." *Review of Economics and Statistics*, 93(3): 1010–1033.

Bhide, Shashanka, and Aasha Kapur Mehta. 2005. "Tackling Poverty Through Panel Data: Rural Poverty in India 1970-1998." *Chronic Poverty Research Centre Working Paper*, 28.

Bignami-Van Assche, Simona, Georges Reniers, and Alexander A Weinreb. 2003. "An Assessment of the KDICP and MDICP Data Quality: Interviewer Effects, Question Reliability and Sample Attrition." *Demographic Research*, 1: 31–76.

Bravo, David, Jere R Behrman, Olivia S Mitchell, Petra E Todd, and Javiera Vásquez. 2008. "Encuesta de Protección Social 2006: Presentación General y Principales Resultados." University of Chile, Department of Economics Working Paper 273.

Bureau of Statistics, Bangladesh. 1995. "Bangladesh Household (Income) Expenditure Survey, BHIES." `www.edi-africa.com/research/khds/introduction.htm/`.

Carolina Population Center, UNC. 2000. "Cebu Longitudinal Health and Nutrition Survey, CPC-UNC Nang Rong." `http://www.cpc.unc.edu/projects/nangrong`.

Carolina Population Center, UNC. 2011. "Cebu Longitudinal Health and Nutrition Survey, CPC-UNC Cebu." `http://www.cpc.unc.edu/projects/cebu`.

Centre for the Study of African Economies, (CSAE). 1995. "The Ethiopian Rural Household Survey, ERHS." `www.edi-africa.com/research/khds/introduction.htm/`.

Falaris, Evangelos M. 2003. "The Effect of Survey Attrition in Longitudinal Surveys: Evidence from Peru, Cote d'Ivoire and Vietnam." *Journal of Development Economics*, 70: 133–157.

Fuwa, Nobuhiko. 2011. "Should We Track Migrant Households When Collecting Household Panel Data? Household Relocation, Economic Mobility, and Attrition Biases in the Rural Philippines." *American Journal of Agricultural Economics*, 93(1): 56–82.

Living Standards Measurement Study, The World Bank. 2004. "Albania Living Standards Measurement Survey, ALSMS." `http://econ.worldbank.org/WBSITE/EXTERNAL/EXTDEC/EXTRESEARCH/EXTLSMS/0,,contentMDK:21369062~pagePK:64168445~piPK:64168309~theSitePK:3358997,00.html`.

Lohana, Hari Ram. 2009. "Poverty Dynamics in Rural Sindh, Pakistan." *Chronic Poverty Research Centre*, Working Paper Number 157.

Magruder, Jeremy R. 2010. "Intergenerational Networks, Unemployment, and Persistent Inequality in South Africa." *American Economic Journal: Applied Economics*, 2(1): 62–85.

Maluccio, John A. 2004. "Using Quality of Interview Information to Assess Nonrandom Attrition Bias in Developing-Country Panel Data." *Review of Development Economics*, 8(1): 91–109.

Norris, Shane A, Linda M Richter, and Stella A Fleetwood. 2007. "Panel Studies in Developing Countries: Case Analysis of Sample Attrition Over the Past 16 Years Within the Birth to Twenty Cohort in Johannesburg, South Africa." *Journal of International Development*, 19(8): 1143–1150.

Outes-Leon, Ingo, and Stefan Dercon. 2008. "Survey Attrition and Attrition Bias in Young Lives." *Young Lives Technical Note*, No. 5.

Popkin, Barry M, Shufa Du, Fengying Zhai, and Bing Zhang. 2010. "Cohort Profile: The China Health and Nutrition Survey: Monitoring and Understanding Socio-Economic and Health Change in China, 1989–2011." *International Journal of Epidemiology*, 39(6): 1435–1440.

RAND Family Life Surveys, (FLS). 1989. "Malaysian Family Life Surveys, MFLS." `http://www.rand.org/labor/FLS/MFLS.html`.

Richter, Linda M, Cesar G Victora, Pedro C Hallal, Linda S Adair, Santosh K Bhargava, Caroline HD Fall, Nanette Lee, Reynaldo Martorell, Shane A Norris, Harshpal S Sachdev, et al. 2012. "Cohort Profile: the Consortium of

Health-Orientated Research in Transitioning Societies." *International Journal of Epidemiology*, 41(3): 621–626.

Rosenzweig, Mark R. 2003. "Payoffs from Panels in Low-Income Countries: Economic Development and Economic Mobility." *American Economic Review*, 93(2): 112–117.

Thomas, Duncan, Firman Witoelar, Elizabeth Frankenberg, Bondan Sikoki, John Strauss, Cecep Sumantri, and Wayan Suriastini. 2012. "Cutting the Costs of Attrition: Results from the Indonesia Family Life Survey." *Journal of Development Economics*, 98(1): 108–123.

Velasquez, Andrea, Maria E Genoni, Luis Rubalcava, Graciela Teruel, and Duncan Thomas. 2010. "Attrition in Longitudinal Surveys: Evidence from the Mexican Family Life Survey." *Northeast Universities Development Consortium Conference (Proceedings).*