

DISCUSSION PAPER SERIES

IZA DP No. 10711

**Attrition in Randomized Control Trials:
Using Tracking Information to Correct Bias**

Teresa Molina Millan
Karen Macours

APRIL 2017

DISCUSSION PAPER SERIES

IZA DP No. 10711

Attrition in Randomized Control Trials: Using Tracking Information to Correct Bias

Teresa Molina Millan

*Nova School of Business and Economics
and IZA*

Karen Macours

Paris School of Economics and INRA

APRIL 2017

Any opinions expressed in this paper are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but IZA takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The IZA Institute of Labor Economics is an independent economic research institute that conducts research in labor economics and offers evidence-based policy advice on labor market issues. Supported by the Deutsche Post Foundation, IZA runs the world's largest network of economists, whose research aims to provide answers to the global labor market challenges of our time. Our key objective is to build bridges between academic research, policymakers and society.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

Attrition in Randomized Control Trials: Using Tracking Information to Correct Bias*

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 labour 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 35 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.

JEL Classification: O1, C93, C52

Keywords: survey non response, sample selectivity, randomized controlled trial, inverse probability weights

Corresponding author:

Teresa Molina Millán
NOVA School of Business and Economics
Universidade Nova de Lisboa
Campus de Campolide
1099-032 Lisboa
Portugal
E-mail: teresa.molina@novasbe.pt

* This research would not have been possible without the support of Ferdinando Regalia of the Inter-American Development Bank (IDB), and the leading role of 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), and the National Science Foundation (SES 11239945 and 1123993). 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 Weerd as well as comments received during presentations at Paris School of Economics and LACEA. All remaining errors and omissions are our own.

1 Introduction

Reliable longitudinal surveys and panel datasets are indispensable tools for the study of economic and demographic dynamics in developing and developed countries. Given the mobility of many target populations, keeping a panel dataset representative of the study population often requires tracking, which can be costly. Data collection cost-concerns are often weighted against the potential of selection bias due to attrition. Most previous studies analyzing the potential biases resulting from missing individuals across waves mainly focus on developed countries.¹ The increased availability of panel data sets in developing countries, often collected specifically for the purpose of impact evaluations, raises the relevance of understanding the implications of high attrition rates. As attrition is often driven by migration in such settings, and as the decision to migrate or the correlates of migration might well be affected by the treatment itself, the challenges posed by attrition could be different than those in developed countries where it is often related to refusals. 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, for instance if the migration decision is related to the fact that treatment effects are different for attriters 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, and omitting migrants can hence lead to an overestimate of the Intent-to-treat (ITT) effects. While a number of longitudinal studies 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 ITT estimations in impact evaluations are not always fully taken into account.

Relatively few panel surveys in developing countries undertake intensive tracking processes to limit migration-related non-response. In a review of tracking protocols on longitudinal surveys we find that most survey designs do not include protocols to track migrants outside their village of origin, sometimes resulting in high attrition rates. This is also the case in many panel surveys collected to estimate impacts of Randomized Control Trials (RCT). As possibilities for longer-term follow-ups of RCTs are increasing, the challenges of tracking and limiting attrition are likely to become even more relevant. 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 started from a review of how RCT studies in development

¹For references see issue number 33 of the *Journal of Human Resources* (1998) on “Attrition in Longitudinal surveys”

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 1 to 60 percent. Notably, the consideration of the potential attrition bias is often limited, in striking contrast with the care given ex-ante to assure random program placement. One fourth of studies do not go beyond testing whether attrition rates between treatment arms are different, and 15 percent of studies do not even show such test. For studies that address attrition in more detail, there also appears to be no standardized approach. Among those studies in which the authors identified non-random attrition, only one-third apply a sample-selection correction method to correct for attrition. On the other hand, about one-third of studies showing balanced attrition rates still apply a sample-selection correction method. Overall, non-parametric bounds (as presented in [Horowitz and Manski \(2000\)](#); [Lee \(2002\)](#); [Kling, Liebman and Katz \(2007\)](#)) and IPW are the most common methodologies applied.

To quantify the implications of different approaches to attrition for a specific case, 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 communities included in the original survey sample, and an *Intense 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 longitudinal and short-panel 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 long-term 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. As discussed, many studies take such a result as enough justification for no further analysis of attrition. 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

²See Appendix [F](#) for details on the selection of papers and the different findings.

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. ³ The ITT estimates suggest that 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 intense 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 to 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 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. Thus, we propose a new method to correct for attrition bias exploiting the similarities, in observables characteristics, between attriters and the intense tracking sample. We build on [Fitzgerald, Gottschalk and Moffitt \(1998\)](#) and [Wooldridge \(2002b\)](#) 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 weight 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 CTP sample. We further show that the observed characteristics have more statistical power in the ITP sample than in the CT sample. Estimates with the new IPW lead to smaller point estimates than the ITT, which are more robust to different specifications of the control. In contrast, applying standard IPW to the full sample leads to less robust estimates, while standard IPW on the regular tracking sample does not appear to correct attrition bias as it leads to point estimates that are similar to the ITT on the regular tracking sample.

We also apply other common approaches to account for attrition and use the information

³The 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 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.

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. This suggests that the monotonicity assumption for the Lee Bounds does not hold in this context, an interpretation in line with the analysis of the correlates of attrition.

Finally, we analyze whether we can rely on proxy information reported by non-migrant household members to get non-attrition biased estimates. We restrict the sample to individuals from whom we have data reported at the communities of origin (proxy information) together with self-reported data. We find that using proxy information on attritors leads to lower treatment effects, although the estimates using proxy information would still have overestimated the treatment effect.

This paper builds on previous work studying attrition bias in household surveys in developing countries. 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 Weerd and Dercon (2011); Velasquez et al. (2010); Thomas et al. (2012)).⁴ Overall, these studies agree on the fact that attritors differ from those who are found in observables characteristics. Even more, attritors may differ among them by destination, as it is the case when we analyze migrants heterogeneity (Beegle, De Weerd and Dercon (2011); Thomas et al. (2012)). That said, Alderman et al. (2001), Falaris (2003), Fuwa (2011) show that estimates are not necessarily biased even if attritors are different from stayers. The Kwazulu-Natal Income Dynamics Study (KIDS) provides a good example to analyze how attrition depends on the outcome of interest. While Alderman et al. (2001) did not find any impact of attrition bias on anthropometric indicators, Maluccio (2004) found evidence of attrition bias on expenditures model using the same database and used survey design characteristics to correct for attrition with a Heckman selection model Maluccio (2004).

To our knowledge, there is only one other paper specifically studying attrition and tracking protocols in the context of a RCT study (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

⁴For instance, Beegle, De Weerd 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.

bias.

The next section discusses different strategies found in the literature to deal with attrition before outlining the rest of the paper.

2 Strategies for dealing with attrition

Attrition can jeopardize the internal and external validity if respondents differ from those who drop-out. Changes in life, including interventions, may affect the decision to move and therefore the probability to attrit. In RCTs attrition may lead to biased estimates if it is correlated with the intervention, i.e. if certain treatment groups have more attrition than others. And even when attrition is balanced between treatment groups and baseline characteristics for non-attriters are balanced, non-random attrition can be a threat to the external validity of the results if the intervention has heterogeneous effects on individuals who are more or less likely to attrit.

To review the potential implications of non-random attrition and the strategies to address them, let's consider a canonical, two-period ($t = 0, 1$), selection model following [Fitzgerald, Gottschalk and Moffitt \(1998\)](#). The outcome variable y_{i1} is regressed on assignment to treatment (T_i) and a vector of variables observed at baseline,

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

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 attriters and non-attriters at time 0 (at baseline). The outcome of interest y_{i1} is observed if $A_{i1} = 0$ and missing due to attrition otherwise. 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 variables (z_{i0}) affecting sample attrition but which are not part of the model of interest.

$$A_{i1}^* = \delta_0 + \delta_T T_i + \delta_1 x_{i0} + \delta_2 z_{i0} + v_{i1} \quad (2)$$

and,

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

If there is correlation between both error terms, ϵ_{i1} and v_{i1} , then estimating equation 1 ignoring equation 2 leads to biased estimates of β . The conditional mean of y_{i1} in the observed sample can be written as

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

To reduce the correlation between both errors, we can correct for sample selection ex-ante, by limiting the number of attritors during the data collection process, or ex-post correcting for attrition bias using parametric and non-parametric econometric techniques.

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

In contexts with high levels of mobility, as often found in developing countries, not tracking migrants can lead to high attrition rates. Tracking 100 percent of an intervention's target population, the strategy followed by the Nicaraguan CCT evaluation, may substantially lower attrition rates and provide data on a sample of migrants similar to respondents that ultimately cannot be found. But it can imply high costs in terms of resources and time. Only few longitudinal data sets track respondents who have moved out their locality of origin. But some successful exceptions show it is 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). 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 of the panel survivor. For summary statistics of attrition rates by tracking protocols see Tables F5-F6 in Appendix F.⁵ Tracking beyond the local level does not seem to be commonly done in impact evaluation panel surveys. Moreover, many studies do not include much information on the tracking protocols, making it hard to quantify how often different practices are used. Some longer-term studies do follow a regular tracking phase with intense tracking of a random subsample of those not found during the RTP. This design was implemented in the 2002 follow-up survey of the U.S Moving to Opportunity (MTO) program (Orr et al., 2003) and has also been used for impact evaluations surveys in developing countries by Baird, Hamory and Miguel (2008); Kremer, Miguel and Thornton (2009); Blattman, Fiala and Martinez (2014). 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.

⁵Among 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).

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 Rya, 2012). The outcome of interest is constructed using observed information on individuals surveyed (y_{i1}) and information reported by others for the sample of attritors (y_{i1}^{proxy}). Hence, we estimate,

$$y'_{i1} = \alpha + \beta T_i + \gamma x_{i0} + \epsilon_{i1} \quad (4)$$

where,

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

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 analyze proxy information for attritors in the Bangladesh Nutrition Survey (1981 to 2000) and concluded that information reported by others, especially on schooling outcomes, was reliable. Drawing on this finding, he uses the proxy reports for other household members who migrated outside the village. This of course relies on the assumption that reliability using information reported by other household members living in the same village, is relevant for proxy reports on far away migrants, for whom outcomes may be harder to observe by prior household members.

2.2 Dealing with attrition *ex-post*

Even after intensive tracking some attrition will almost always remain, which may 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 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, 2002a), 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, $\epsilon_{i1}|x_{i0}$ is independent of v_{i1} but $\epsilon_{i1}|x_{i0}$ is not independent of z_{i0} , that is selection bias arises for a vector of observed characteristics (z_{i0}). On the other hand, if $\epsilon_{i1}|x_{i0}$ is independent from z_{i0} but $\epsilon_{i1}|x_{i0}$ is not independent from v_{i1} , selection on unobservables complicates the identification.

Under the assumption of selection based on observables unbiased estimates can be obtained using weighted least square regression (Fitzgerald, Gottschalk and Moffitt, 1998; Wooldridge, 2002a). To model the probability of sample selection on observables Fitzgerald, Gottschalk and

Moffitt (1998) make the following assumption on Equations 1-2,

Assumption 1.

1. y_{i1} is observed whenever $A_{i1} = 0$
2. A_{i1} , z_{i0} and x_{i0} are always observed for all i
3. v_{i1} is independent of $\epsilon_{i1}|x_{i0}, v_{i1}|z_{i0}, x_{i0} \sim \text{Normal}(0, 1)$
4. For all $z \in Z$, $x \in X$, $P(A_{i1} = 0|z_{i0}, x_{i0}) > 0$

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

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

Given 1 and 2 and under standard regularity conditions we can use predicted probabilities of being surveyed to correct for non-random selection on observables (Wooldridge, 2002a). The standard procedure to construct IPW consists on estimating the probability of being surveyed, conditional on a specified set of covariates, using the complete target population. Behrman, Parker and Todd (2009), for instance, apply IPW when estimating medium-term impacts of the Oportunidades/Progresa CCT program in Mexico. As the correlates of attrition are significantly different between treatment arms, they estimate separate weights for each of the experimental group. Applying these weights adjust for the differences in baseline characteristics between treatment arms that arise because of attrition.

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.⁶

As both IPW and Heckman’s correction selection model are based on strong assumptions, it has become relatively common to use non-parametric techniques to determine intervals of estimates size.

⁶This implies that ideally interviewers should be randomly assigned. We found only one study in which a Heckman Selection Model was used to correct for attrition using information on a randomized survey design (Dinkelman and Martínez A, 2014).

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 have 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. Recent papers following this approach include- [Karlan and Valdivia \(2011\)](#); [Blattman, Fiala and Martinez \(2014\)](#); [Drexler, Fischer and Schoar \(2014\)](#).

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 ([Kremer, Miguel and Thornton \(2009\)](#)); [Baird, McIntosh and Özler \(2011\)](#)); [Hidrobo et al. \(2014\)](#)); [Cunha \(2014\)](#)); [Drexler, Fischer and Schoar \(2014\)](#)).

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.

In this paper we do not implement a Heckman selection model to correct for attrition because of the lack of credible exclusion restrictions. Instead, we propose an alternative approach that builds on [Fitzgerald, Gottschalk and Moffitt \(1998\)](#) methodology of constructing a

model specific IPW, and exploit similarities between difficult-to-find respondents and attritors to correct for attrition bias.⁷ To motivate the approach we use data from the long-term evaluation of a randomized CCT program in Nicaragua. The next section introduces the program, the evaluation design and the data collection. 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. The last section presents insights of the cost of conducting ITP in terms of enumerators days of work and concludes.

3 Red de Proteccion Social: Program design, Evaluation and Data⁸

3.1 Program design and Evaluation

The Red de Proteccion Social 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 *Progresas/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 a 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 expo-

⁷This approach has similarities with [Behaghel et al. \(2015\)](#) selectivity correction procedure, as in order to correct for non-random attrition both methods use information on those who were difficult to find.

⁸See [Flores and Maluccio \(2005\)](#) for additional details on the program and the experimental design.

sure in the early treatment localities than in the late treatment localities due to the eligibility criteria for the education transfer and pre-program school drop out 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.⁹

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 1,756 households randomly selected for interview during the year 2000 baseline surveys in the early and late treatment areas, as well as an 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 2000.¹⁰ 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 communities reduced attrition in the sample due to seasonal migration.

3.3 Tracking

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 communities of origin and some migrants were followed to other communities 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.

⁹Off-farm employment is measured as a dichotomous variables that takes value one if the individual is economically active (in wage or self-employment) outside of the family farm, and zero otherwise.

¹⁰To keep the sample representative of the target population, all estimates include sample weights constructed at the locality level.

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 communities 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 question 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 of 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.

3.4 Survey Attrition

Final attrition rates for males in the cohort of interest are low for a 10 years 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 10.17 percent (Table 1).¹¹

Table 2 shows attrition 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.¹² After RTP attrition rates were still relatively large, 26 percent, but differences between treatment arms are not significantly different from zero. The last row of Table 2 shows the response rates after conducting ITP conditional of not being found during

¹¹Attrition includes those who have migrated and those who refused to be interviewed, which account for less than 0.1 percent of those non-respondents.

¹²That 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.

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.

Table A1 in Appendix A shows how attrition affected balance of variables observed in the baseline census. Balance results are shown for the target sample of 1,138 boys in the first column, for the sample of 1,022 boys tracked in the second column and the sample of 841 boys tracked after the regular tracking phase in the third column.¹³ The table shows that the randomization resulted in very few significant differences between the early and late treatment group, as expected. After regular tracking, a number of additional baseline variables were off balance, in particular related to parental education and household demographics. 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 those 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:

$$Y_{i2010} = \alpha + \beta T_i + \gamma X_{i2000} + \epsilon_i \quad (5)$$

where Y_{2010} is the outcome of interest in 2010, T is an ITT indicator that takes value of one for children in communities 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).¹⁴ The first four columns show the results with the complete

¹³The table and all subsequent analyses do not include information for three boys that were tracked at follow-up but for whom data on education at baseline is missing.

¹⁴The sample used includes 1,006 individuals found and from whom we do have information on grades attained

tracked sample under different specifications for X: only strata fixed effects (column 1); strata and three monthly age fixed effects and baseline education (column 2) ¹⁵; the same controls plus a vector of the covariates that were off-balance after the relevant tracking phase (column 3) ¹⁶; and column 4 in addition has 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. ¹⁷

Estimates on the full sample show that boys coming from communities randomly assigned to early treatment have 0.593 more grades attained than boys from the late treatment group. Including 3 monthly age fixed effects and controlling for baseline level of education in the regression decreases the point estimates on assignment to early treatment to 0.427 grades attained. Adding the variables that were not balanced at baseline decreases the point estimate to 0.362. The three point estimates are significantly different from zero at the 5 percent level. Finally, including 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 four 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 0.865 and 0.613 in the first and second specification respectively, both significantly different from zero at the 5 and 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 decreases the point estimate to 0.399, that is, an adjustment of 35 percent with respect to the estimate in the 2nd specification. The point estimate falls further after including other baseline controls to 0.326. The last two rows report p-values for testing the equality of coefficients at different stages during the tracking process. The ITT estimate after RTP is 45 percent larger, and significantly different from the final estimate using the CTP sample, in the first two specifications. Differences between CTP and RTP are smaller and not significant with the expanded set of controls.

and off-farm employment in 2010. The sample does not include 15 deceased individuals found in 2010.

¹⁵A set of dummies indicating whether the individual had 1,2,3 or at least 4 years of education at baseline.

¹⁶After complete tracking the off-balance controls include baseline controls for whether the individual was working, for the number of villagers with family ties, for the village population size and for a productive asset index (see Appendix D). After regular tracking the off-balance controls include as well controls for whether the mother had no education, for whether the mother had at least three years of education, for whether the individual is son of the household head, for the number of children of the household head and for whether the head is female. In the regression on the sample targeted during ITP the off-balance controls include baseline controls for the number of villagers with family ties, for the village population size and for a productive asset index. The probability of attrition prior to program started at the *comarca* level is not included as control (even if it is off-balance) as it captures attrition before the program started but after the participants knew the treatment group they were assigned to.

¹⁷The four specifications we present reflect 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.

The last four 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 45 percent 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 intense tracking phase.¹⁸ The figure shows that individuals found during intensive tracking in close-by regions and Managua are driving point estimates down. It also shows 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

Even if response rates are balanced by treatment group, results in Section 4 confirm the importance of understanding the correlates of attrition to make informed assumptions about the

¹⁸These 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.

nature of selection into the final sample. 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. 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.¹⁹

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 behaviour, 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 observables 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 the appendix.

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

¹⁹Migration 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 potentially could be more relevant for households with complementary productive assets.

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 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 a large share moved to locations in different municipalities and departments (see Table 5). Note also that the 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).

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.²⁰

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 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 pop-

²⁰Using 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 Weerd and Dercon (2011) find similar patterns in their analysis on economic mobility in Tanzania.

ulation in the presence of heterogeneous treatment effects. Moreover such heterogeneity seems likely given that some of the predictors of attrition are the same as the predictors of compliance. We therefore calculate weights to correct for selective attrition.

5.2 Obtaining Probability Weights using information from tracking

The intense tracking strategy provides valuable information to calculate weights for the attrition selection correction. [Fitzgerald, Gottschalk and Moffitt \(1998\)](#) and [Wooldridge \(2002b\)](#) have shown that consistent estimates can be obtained using weighted least squares regression, under the basic assumption that the probability of being found is driven by observable characteristics. We use a modified version of this standard inverse probability weighting adjustment and overweight 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. The basic assumption underlying our estimation strategy is that individuals that were hard to find are also more similar in unobservables to those that ultimately were not found.

We modify the IPW estimates by exploiting these similarities, which allows weakening the assumptions underlying the attrition weights. In particular, we construct weights using only the individuals tracked during the intensive tracking phase.

If those found during the intense tracking phase are more similar to attritors, assumption 1 and 2 in [Section 2.2](#) might be more plausible if selection is modeled for the intensive tracking only.

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

$$\hat{p}_{i1}^{ITP} = P(z_{i0}, x_{i0}, \hat{\delta}_1^{ITP}, \hat{\delta}_2^{ITP}) \quad (6)$$

Respondents interviewed during the RTP are assigned a $P(A_{i1} = 0 | z_{i0}, x_{i0}) = 1$. We obtain the target population density function by weighting the conditional density in the ITP sample using $w_{i1}^{ITP}(z, x) = \left[\frac{1}{\hat{p}_{i1}^{ITP}} \right]$ and weighting the conditional density in the RTP sample using $w_{i0}^{RT}(z, x) = 1$. 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.

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 endogenous variables that may be driving selection (henceforth, Z) 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 . Hence as potential predictors we consider 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. We first estimate bivariate regressions in which each potential 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 *comarca* level. The testing was conducted separately for the early and late treatment group.²¹ Results can be found in Tables B1, B3 and B4 in Appendix A. 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 group 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. Each two columns of Table 6 reports 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. Table 6 confirms 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.

²¹We include estimates of early and late treatment *comarcas* in the same regression but interact the variable of interest with the treatment status to obtain separate estimates for early and late treatment group.

The bottom panel of Table 6 shows how the power of the different set of variables differs by tracking phase. 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 community ex-ante attrition rates). 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 in Table 6 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 7 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 four 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. We compare these estimates with the OLS estimates in the first 4 columns of Table 3. We also compare with WLS estimates using weights estimated with the entire sample (columns 5 to 8). In the bottom of each panel, we show WLS estimates only using the sample from the regular tracking.

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 two specifications. In the models in which we do not control for any baseline characteristics the new WLS is 0.361, adding 3 monthly dummies and the set of indicators for baseline education increase precision but does not affect the value of the point estimate. Adding off balance covariates and other baseline controls (columns 3 and 4) reduces the point estimate slightly to 0.344 and 0.317 respectively, with both coefficients 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 four specifications (columns 5 to 8). The WLS estimates appear as sensitive as the OLS estimates to the inclusion of different set of controls, suggesting that the standard weights by themselves may not correct for all the selective attrition.

The education estimates after the regular tracking, are even more sensitive to changes in the set of control variables, 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.²² Overall, the attrition correction from adding the controls and estimating the new IPW is substantial. That said, the biggest difference in estimates is

²²Results 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.

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 CT 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 and off-balance controls in the specification leads to higher point estimates. This confirms the earlier finding that estimates after RTP appear to overestimate the treatment effects in this sample. The lower compliance among individuals 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 attriters. 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.²³

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 communities and the maximum (*minimum*) value of non-attrited distribution to the attrited from late treatment communities. 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 attriters from early treatment communities and the mean plus (*minus*) 0.75 standard deviations of the non-attrited late treatment distribution to the attriters from late treatment communities. 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 \(2009\)](#)

²³Hence estimates are to be compared with the 2nd specification in Table 3 and 7. Estimates for the other specifications can be found in appendix Tables A2-A3.

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

Table 8 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 non-informative 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 observables characteristics, but they do not correct for any heterogeneity in the not-found sample. Section 3.4 suggests that in our sample we do have selection on observables 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 do no longer belong to the baseline household. Following [Rosenzweig \(2003\)](#) we compare self-report and proxy reports to analyze data reliability. 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 A4 and A5 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 9 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). Column 3 and 4 reports 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 4). 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 0.1.²⁴ 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.

Column 5 to 6 extends 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 attritors to estimate Equation 4 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

²⁴Estimates with full set of baseline controls are qualitative similar, but differences are smaller.

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 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 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. While this is well known, many studies based on RCTs appear to pay limited attention to attrition, and most do not go beyond reporting that attrition rates are balanced between experimental groups. 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 labour market outcomes. Sample selection is driven by observables characteristics and those found during an intensive tracking phase are more similar to those not found, a result that mirror 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 intense 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 involved in tracking which are indeed non-ignorable. 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, it is important to 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 sample sizes are large enough. Using the information from hard-to-find individuals to calculate attrition-weights allows getting more mileage out of such intensive tracking, and as such can help improve the trade-off between costs and potential bias.

References

- 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. Dufto (eds.) Handbook of Economic Field Experiments*, Volume 1. Elsevier.
- 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, 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.
- 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 Weerd, 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.
- 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.
- Cunha, Jesse M.** 2014. "Testing Paternalism: Cash Versus In-Kind Transfers." *American Economic Journal: Applied Economics*, 6(2): 195–230.
- Deaton, Angus, and Nancy Cartwright.** 2016. "Understanding and Misunderstanding Randomized Controlled Trials." *NBER Working paper*, No.w22595.
- 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>.
- 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, Rema Hanna, and Stephen P Rya.** 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 Analysis of Sample Attrition in Panel Data: The Michigan Panel Study of Income Dynamics.” *Journal of Human Resources*, 33(2): 251–299.
- 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.
- 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.
- Hill, Daniel H, and Robert J Willis.** 2001. “Reducing Panel Attrition: A Search for Effective Policy Instruments.” *Journal of Human Resources*, 416–438.
- 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.

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

- 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.** 2002a. *Econometric Analysis of Cross Section and Panel Data*. The MIT press.
- Wooldridge, Jeffrey M.** 2002b. "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		Diff means (s.e)
	Early Treatment	Late Treatment	
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)

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

Note: Attrition rates remain balanced if we include those who were found dead as attritors.

Table 3: Intent-To-Treat

	Complete Tracking Phase				Regular Tracking Phase				Intense 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** (0.29)	0.427** (0.17)	0.362** (0.18)	0.319* (0.16)	0.865** (0.33)	0.613*** (0.20)	0.451** (0.20)	0.350* (0.18)	-0.745 (0.58)	-0.481 (0.37)	-0.400 (0.43)	0.053 (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
<i>Comparing coefficients at different stages during the tracking process: χ^2</i>												
P-val: $T^{CTP} - T^{TP} = 0$					0.0330	0.0361	0.3202	0.7219	0.0241	0.0149	0.0598	0.5342
P-val: $T^{RTP} - T^{TP} = 0$									0.0231	0.0143	0.0689	0.5417
Off-Farm Employment												
Early-Treatment	0.059* (0.03)	0.061* (0.03)	0.056* (0.03)	0.055** (0.02)	0.087** (0.03)	0.088** (0.03)	0.081** (0.03)	0.083*** (0.03)	-0.071 (0.05)	-0.097* (0.06)	-0.097* (0.06)	-0.125** (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
<i>Comparing coefficients at different stages during the tracking process: χ^2</i>												
P-val: $T^{CTP} - T^{TP} = 0$					0.0041	0.0076	0.0156	0.0054	0.0093	0.0030	0.0028	0.001
P-val: $T^{RTP} - T^{TP} = 0$									0.0073	0.0027	0.0025	0.001
Obs.	1006	1006	1006	1006	827	827	827	827	179	179	179	179

Standard errors are clustered at locality level. Robust s.e in brackets.*** p<0.01, ** p<0.05, * p<0.1

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 controls include baseline controls for whether the individual was working, for the number of individuals with family ties in village, for the village population size and for a productive asset index (2nd principal component of household wealth). After regular tracking the off-balance controls include as well controls for whether the mother had no education, for whether the mother had at least three years of education, for whether the individual is son of the household head, for the number of children of the household head and for whether the head is female. The regression on the sample targeted during ITP the off-balance controls include baseline controls for the number of individuals with family ties in village, for the village population size and for 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.

Table 4: Baseline Characteristics (2000) by tracking phase.

	Complete Track. (1)	Regular Track. (2)	Intensive Track. (3)	Attritors (4)	Diff. (1)-(4)	Diff. (3)-(4)
<i>Proxy's of permanent residence in village</i>						
Own house (= 1)	0.844	0.861	0.764	0.709	0.135* (0.068)	0.055 (0.092)
House is obtained in exchange for serv./labor (= 1)	0.069	0.047	0.165	0.167	-0.098 (0.059)	-0.001 (0.082)
Address in hacienda (= 1)	0.134	0.115	0.219	0.217	-0.083 (0.064)	0.003 (0.087)
Address in hacienda & house rented (= 1)	0.063	0.048	0.130	0.152	-0.089 (0.054)	-0.022 (0.088)
<i>Social Capital</i>						
Family network size (individuals)	85.328	88.544	70.901	44.078	41.250*** (6.807)	26.823*** (8.008)
Population size village	485.121	481.333	502.115	514.407	-29.286 (48.301)	-12.292 (58.114)
<i>Village Characteristics</i>						
Village affected by hurrican Mitch (= 1)	0.871	0.877	0.847	0.902	-0.031 (0.032)	-0.055 (0.044)
Altitude of village	628.461	630.422	619.662	635.981	-7.521 (28.461)	-16.320 (33.410)
Village in coffee producing region (= 1)	0.783	0.784	0.777	0.787	-0.005 (0.053)	-0.010 (0.068)
Distance to night light (meters)	17096.88	16568.36	19467.58	21805.35	-4708.47*** (1141.06)	-2337.78** (1009.57)
Live in Tuma region (= 1)	0.410	0.388	0.506	0.728	-0.318*** (0.047)	-0.222*** (0.054)
Live in Madriz region (= 1)	0.203	0.213	0.161	0.064	0.140** (0.052)	0.097** (0.042)
<i>Household Characteristics: Economic activities & Assets</i>						
Household head main occupation is ag. (= 1)	0.823	0.834	0.774	0.733	0.089 (0.057)	0.040 (0.084)
Size of landholdings ('000 sq meters)	16.260	16.749	14.066	18.584	-2.325 (5.740)	-4.518 (6.629)
Number of parcels of land	0.964	0.979	0.899	0.711	0.253*** (0.075)	0.189 (0.123)
Log predicted expenditures (pc)	7.727	7.737	7.679	7.768	-0.041 (0.052)	-0.089 (0.063)
Wealth index - housing characteristics	0.048	0.061	-0.007	0.369	-0.321 (0.268)	-0.377 (0.292)
Wealth index - productive assets	-0.005	0.028	-0.150	-0.421	0.416*** (0.121)	0.271* (0.150)
Wealth index - other assets	-0.038	-0.060	0.061	0.263	-0.301** (0.116)	-0.202 (0.126)

Table 4: Baseline Characteristics (2000) by tracking phase.(Cont.)

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

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

Table 5: Compliance, Baseline Attrition and Migration (2000) by tracking phase.

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

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

Table 6: Linear Probability Model. Correlates of Attrition.

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

Table 6: LPM. Determinants of Attrition. Boys 9-12 years old (Continue)

	CTP Sample		RTP Sample		ITP Sample	
	(1)	ET Interaction (2)	(3)	ET Interaction (4)	(5)	ET Interaction (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						
<i>Covariates and interaction terms together</i>						
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	
<i>Covariates and interaction terms separately</i>						
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

Standard errors are clustered at locality level. Robust s.e in parenthesis. *** p<0.01, ** p<0.05, * p<0.1

Table 7: 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 (0.32)	0.363* (0.18)	0.344* (0.18)	0.317* (0.16)	0.503 (0.30)	0.392** (0.19)	0.329* (0.19)	0.282* (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** (0.32)	0.655*** (0.21)	0.531** (0.21)	0.362** (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** (0.03)	0.061* (0.03)	0.055* (0.03)	0.054** (0.02)	0.060** (0.03)	0.063** (0.03)	0.058* (0.03)	0.055** (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** (0.03)	0.086** (0.04)	0.075** (0.03)	0.076*** (0.03)
R-squared					0.03	0.05	0.08	0.11
Obs.					827	827	827	827

Standard errors are clustered at locality level. Robust s.e in parenthesis.*** p<0.01, ** p<0.05, * p<0.1

Notes: 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 controls include baseline controls for whether the individual was working, for the number of individuals with family ties in village, for the village population size and for a productive asset index (2nd principal component of household wealth). After regular tracking the off-balance controls include as well controls for whether the mother had no education, for whether the mother had at least 3 years of education, for whether the individual is son of the household head, for the number of children of the household head and for whether the head is female. The regression on the sample targeted during ITP the off-balance controls include baseline controls for the number of individuals with family ties in village, for the village population size and for 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.

Table 8: 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 <i>à la</i> Kling						
OLS	0.427** (0.17)			0.613*** (0.20)		
Worst-Best Scenario		-1.181*** (0.27)	2.132*** (0.26)		-3.349*** (0.30)	4.685*** (0.34)
Mean \pm 0.75 s.d		0.001 (0.16)	1.029*** (0.16)		-0.425** (0.18)	2.066*** (0.20)
Mean \pm 0.50 s.d		0.172 (0.16)	0.858*** (0.15)		-0.010 (0.17)	1.651*** (0.19)
Mean \pm 0.25 s.d		0.344** (0.15)	0.686*** (0.15)		0.405** (0.17)	1.236*** (0.18)
Case 3: Lee Bounds						
Lee Bounds		0.463** [0.234]	0.698*** [0.246]		0.471 [0.36]	1.32*** [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 <i>à la</i> Kling						
OLS	0.061* (0.03)			0.088** (0.03)		
Worst-Best Scenario		-0.060* (0.03)	0.161*** (0.03)		-0.219*** (0.04)	0.316*** (0.03)
Mean \pm 0.75 s.d		0.005 (0.03)	0.119*** (0.03)		-0.050* (0.03)	0.227*** (0.03)
Mean \pm 0.50 s.d		0.024 (0.03)	0.100*** (0.03)		-0.004 (0.03)	0.181*** (0.03)
Mean \pm 0.25 s.d		0.043 (0.03)	0.081*** (0.03)		0.042 (0.03)	0.135*** (0.03)
Case 3: Lee Bounds						
Lee Bounds		0.044 [0.029]	0.062*** [0.023]		0.027 [0.041]	0.101*** [0.027]
Obs.		1122	1122		1122	1122
%-trimmed		0.017	0.017		0.068	0.068

Manski Bounds and Bounds *à la* Kling: Standard errors are clustered at locality level. Robust s.e in parenthesis. *** p<0.01, ** p<0.05, * p<0.1. 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.

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.

Table 9: 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** (0.17)	0.613*** (0.19)	0.523*** (0.19)	0.512*** (0.19)	0.424** (0.17)	0.516*** (0.18)
Outcome Mean	5.78	5.88	5.86	5.84	5.68	5.75
Obs.	1006	826	881	881	1072	934
R-square	0.42	0.43	0.42	0.42	0.41	0.42
Off-Farm Employment						
Early-Treatment	0.061** (0.03)	0.088*** (0.03)	0.086** (0.03)	0.092*** (0.03)	0.058* (0.03)	0.090** (0.03)
Outcome Mean	0.86	0.86	0.87	0.86	0.86	0.85
Obs.	1006	827	881	881	1071	933
R-square	0.04	0.05	0.05	0.05	0.04	0.05

Standard errors are clustered at locality level. Robust s.e in parenthesis.*** p<0.01, ** p<0.05, * p<0.1

Notes: 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.

Figure 1: Cost Analysis: ITT estimates on Grades attained during the ITP.

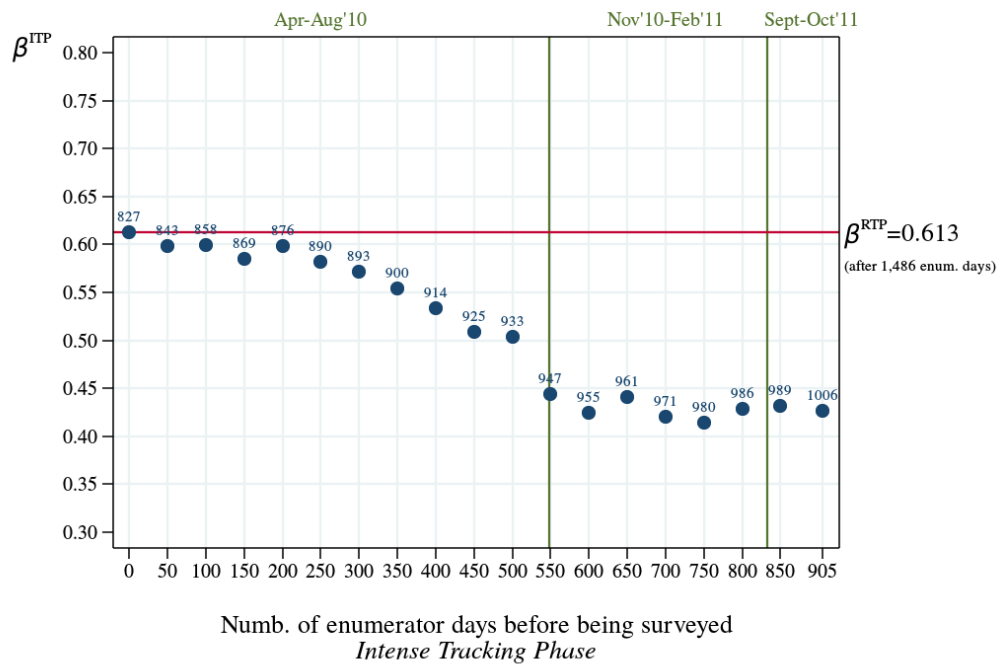
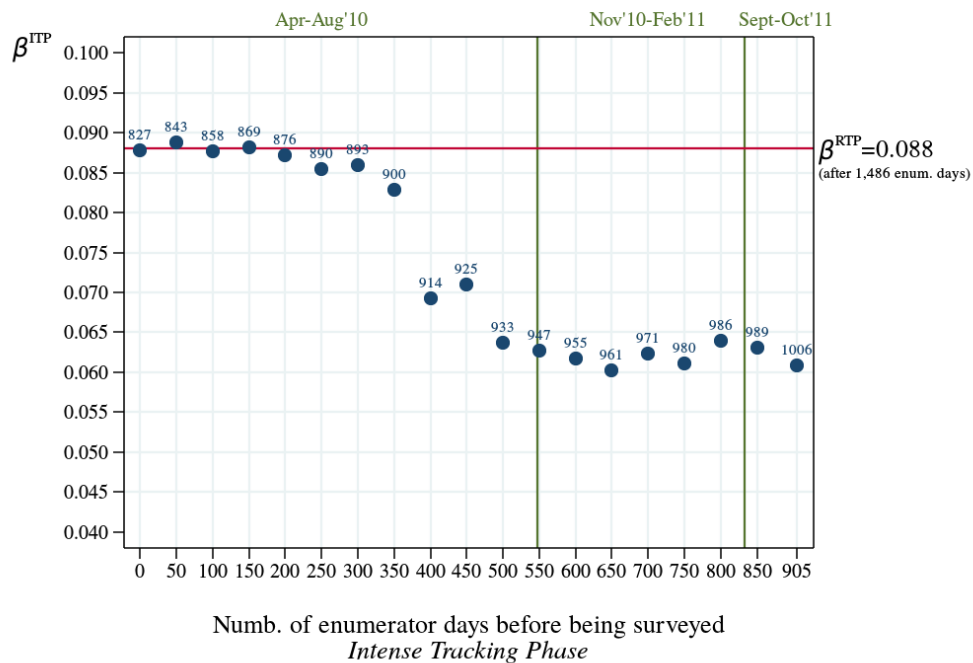


Figure 2: Cost Analysis: ITT estimates on Off-Farm Employment during the ITP.



Notes (Figures 1 & 2): 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. 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”

Teresa Molina Millán and Karen Macours

April 13, 2017

A Appendix

List of Variables

- **Early-Treatment:** An ITT indicator that takes value of one for children in communities 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 hurrican Mitch (= 1):** Dichotomous variable that takes value of one for villages affected by the hurricaine 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 comarca:** Attrition rate at the comarca 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

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)
<i>Variables Indicating Very Early Attrition</i>			
Probability of attrition prior to program start in comarca	-0.078** (0.034)	-0.076** (0.035)	-0.080** (0.035)
Nobody of target sample attrited before program start	0.100 (0.15)	0.096 (0.15)	0.12 (0.15)
<i>Proxy's of permanent residence in village</i>			
Own house (= 1)	-0.052 (0.043)	-0.061 (0.043)	-0.020 (0.046)
House is obtained in exchange for service/labor (= 1)	0.019 (0.039)	0.019 (0.036)	-0.018 (0.032)
Address in hacienda (= 1)	-0.022 (0.058)	-0.0089 (0.056)	-0.028 (0.058)
Address in hacienda & house rented (= 1)	0.032 (0.033)	0.043 (0.031)	0.015 (0.031)
<i>Social Capital</i>			
Family network size (individuals)	23.5** (10.3)	26.0** (10.8)	27.8** (12.1)
Population size village	257.6*** (88.6)	256.0*** (89.3)	247.1*** (95.1)
<i>Village Characteristics</i>			
Village affected by hurrican Mitch (= 1)	-0.062 (0.050)	-0.055 (0.052)	-0.057 (0.051)
Altitude of village	-21.1 (34.6)	-22.1 (36.2)	-25.5 (36.4)
Village in coffee producing region (= 1)	-0.018 (0.065)	-0.0084 (0.066)	-0.022 (0.067)
Distance to night light (meters)	2276.2 (2642.3)	1882.3 (2629.1)	1276.3 (2642.7)
Live in Tuma region (= 1)	0.19 (0.15)	0.19 (0.14)	0.16 (0.14)
Live in Madriz region (= 1)	0.017 (0.12)	0.029 (0.13)	0.048 (0.12)
<i>Household Characteristics: Economic activities & Assets</i>			
Household head main occupation is ag. (= 1)	-0.021 (0.038)	-0.020 (0.040)	-0.0054 (0.046)
Size of landholdings ('000 sq meters)	-1.99 (1.85)	-2.41 (1.79)	-2.73 (2.01)
Number of parcels of land	-0.036 (0.073)	-0.0075 (0.074)	-0.017 (0.075)
Log predicted expenditures (pc)	-0.0089 (0.022)	0.0021 (0.025)	0.027 (0.026)
Wealth index - housing characteristics	0.20 (0.16)	0.22 (0.15)	0.32 (0.16)
Wealth index - productive assets	-0.26** (0.11)	-0.26** (0.11)	-0.21** (0.12)
Wealth index - other assets	-0.040 (0.18)	-0.061 (0.19)	-0.13 (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)
<i>Household Characteristics: Demographics</i>			
Father not living in same household (= 1)	0.017 (0.031)	0.012 (0.027)	0.042 (0.029)
Mother not living in same household (= 1)	0.011 (0.017)	0.013 (0.017)	0.029 (0.018)
Child of household head (= 1)	-0.016 (0.024)	-0.017 (0.024)	-0.060** (0.024)
Number of children of household head	-0.24 (0.22)	-0.32 (0.23)	-0.51** (0.23)
Female household head (= 1)	0.030 (0.020)	0.020 (0.017)	0.037** (0.018)
Age of household head	0.34 (0.87)	0.41 (0.81)	0.52 (0.88)
Number of household members	-0.047 (0.19)	-0.17 (0.22)	-0.21 (0.24)
Nuclear household (= 1)	-0.017 (0.040)	-0.021 (0.043)	-0.027 (0.044)
Multigenerational household (= 1)	-0.037 (0.035)	-0.015 (0.038)	-0.00043 (0.035)
Other household structure (= 1)	0.054* (0.028)	0.036 (0.028)	0.027 (0.031)
Number of children aged 0-8	0.021 (0.11)	-0.098 (0.11)	-0.067 (0.12)
Number of children aged 9-12	-0.036 (0.058)	-0.076 (0.063)	-0.082 (0.071)
<i>Household Characteristics: Education</i>			
Distance to nearest school (minutes)	0.74 (4.76)	0.89 (4.65)	2.92 (4.71)
Mother no grades attained (= 1)	-0.046 (0.037)	-0.052 (0.042)	-0.097** (0.041)
Mother 3 plus grades attained (= 1)	0.072 (0.044)	0.065 (0.042)	0.11*** (0.038)
Household head no grades attained (= 1)	0.0049 (0.031)	0.018 (0.034)	0.0017 (0.039)
Household head 3 plus grades attained (= 1)	0.025 (0.029)	0.033 (0.027)	0.047 (0.031)
<i>Individual Characteristics</i>			
Age at start of transfer in months	-0.046* (0.026)	-0.068** (0.031)	-0.090* (0.046)
Highest grade attained	0.071 (0.16)	0.096 (0.15)	0.14 (0.16)
No grades attained (= 1)	-0.0082 (0.064)	-0.019 (0.063)	-0.029 (0.066)
Worked in last week (= 1)	-0.054* (0.027)	-0.062** (0.029)	-0.063** (0.028)
Participated in some economic activity (= 1)	-0.0086 (0.035)	-0.012 (0.034)	-0.015 (0.036)
Obs.	1138	1022	841

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

Table A2: Lower and Upper Bound. 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** (0.29)			0.865** (0.33)		
Worst-Best Scenario		-1.056*** (0.31)	2.218*** (0.31)		-3.223*** (0.35)	4.753*** (0.35)
Mean \pm 0.75 s.d		0.114 (0.26)	1.130*** (0.27)		-0.319 (0.26)	2.154*** (0.25)
Mean \pm 0.50 s.d		0.283 (0.26)	0.961*** (0.26)		0.093 (0.25)	1.742*** (0.25)
Mean \pm 0.25 s.d		0.453* (0.26)	0.791*** (0.26)		0.505** (0.24)	1.330*** (0.24)
Stratif.,Age and Educ. F.E. +Unbalanced Controls						
OLS	0.426** (0.16)			0.558*** (0.20)		
Worst-Best Scenario		-1.180*** (0.26)	2.166*** (0.26)		-3.355*** (0.31)	4.686*** (0.34)
Mean \pm 0.75 s.d		0.019 (0.16)	1.059*** (0.16)		-0.418** (0.17)	2.076*** (0.19)
Mean \pm 0.50 s.d		0.193 (0.15)	0.885*** (0.15)		-0.003 (0.17)	1.660*** (0.18)
Mean \pm 0.25 s.d		0.366** (0.15)	0.712*** (0.15)		0.413** (0.16)	1.244*** (0.17)
Stratif.,Age and Educ. F.E. +Unbalanced and Baseline Controls						
OLS	0.350** (0.17)			0.424** (0.20)		
Worst-Best Scenario		-1.214*** (0.28)	2.140*** (0.28)		-3.410*** (0.35)	4.591*** (0.35)
Mean \pm 0.75 s.d		-0.009 (0.17)	1.033*** (0.17)		-0.493** (0.20)	1.989*** (0.19)
Mean \pm 0.50 s.d		0.164 (0.16)	0.859*** (0.16)		-0.080 (0.18)	1.575*** (0.17)
Mean \pm 0.25 s.d		0.338** (0.15)	0.685*** (0.15)		0.334* (0.17)	1.161*** (0.17)

Standard errors are clustered at locality level. Robust s.e in parenthesis.*** p<0.01, ** p<0.05, * p<0.1.

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.

Table A3: Lower and Upper Bound. 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 \pm 0.75 s.d		0.003	0.116***		-0.051*	0.223***
		(0.03)	(0.03)		(0.03)	(0.03)
Mean \pm 0.50 s.d		0.022	0.097***		-0.005	0.178***
		(0.03)	(0.03)		(0.03)	(0.03)
Mean \pm 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 \pm 0.75 s.d		0.009	0.124***		-0.046*	0.231***
		(0.03)	(0.03)		(0.03)	(0.03)
Mean \pm 0.50 s.d		0.028	0.105***		0.000	0.185***
		(0.03)	(0.03)		(0.03)	(0.03)
Mean \pm 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 \pm 0.75 s.d		-0.002	0.114***		-0.055*	0.220***
		(0.03)	(0.03)		(0.03)	(0.03)
Mean \pm 0.50 s.d		0.017	0.095***		-0.009	0.174***
		(0.03)	(0.03)		(0.03)	(0.03)
Mean \pm 0.25 s.d		0.037	0.075**		0.037	0.128***
		(0.03)	(0.03)		(0.03)	(0.03)

Standard errors are clustered at locality level. Robust s.e in parenthesis.*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

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.

Table A4: Percentage of respondents from whom we have proxy information on education and labor outcomes.

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

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

Table A5: Percentage of respondents with accurate proxy information and direction of the bias for respondents without accuracy information.

	Complete Tracking Phase		Regular Tracking Phase		Intense 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 (0.07)	-0.075 (0.07)	0.060 (0.07)	0.001 (0.10)	-0.034 (0.14)	-0.128 (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 (0.08)	-0.018 (0.08)	-0.011 (0.11)	-0.121 (0.09)	0.131 (0.12)	0.087 (0.12)
Obs.	175	175	100	100	75	75
Control Mean	0.66	-0.31	0.60	-0.36	0.75	-0.24

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

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 communities.²⁵ 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 have a family network in the communities of origin, they may be harder to track once they have moved on.²⁶ Table 4 and Table B1 confirm that proxy's 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 attriters were also more likely to come from female headed and smaller households and were less likely to live with their biological parents. Attriters 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 significant 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).

²⁵Such 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.

²⁶To 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 *comarca* 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
<i>Variables Indicating Very Early Attrition</i>				
Probability of attrition prior to program start in comarca	-0.323 (0.375)	-0.211 (0.140)	0.781	0.235
Nobody of target sample attrited before program start	0.088** (0.040)	0.093*** (0.026)	0.914	0.001
<i>Proxy's of permanent residence in village</i>				
Own house (= 1)	0.065 (0.065)	0.146** (0.064)	0.381	0.054
House is obtained in exchange for service/labor (= 1)	-0.120 (0.113)	-0.154** (0.071)	0.798	0.063
Address in hacienda (= 1)	-0.035 (0.068)	-0.111 (0.072)	0.446	0.277
Address in hacienda & house rented (= 1)	-0.059 (0.082)	-0.289** (0.136)	0.153	0.091
<i>Social Capital</i>				
Family network size (individuals)	0.001*** (0.000)	0.001*** (0.000)	0.733	0.000
Population size village	-0.000 (0.000)	-0.000 (0.000)	0.948	0.844
<i>Village Characteristics</i>				
Village affected by hurrigan Mitch (= 1)	-0.002 (0.042)	-0.096*** (0.024)	0.059	0.001
Altitude of village	-0.000 (0.000)	0.000 (0.000)	0.620	0.856
Village in coffee producing region (= 1)	-0.002 (0.048)	-0.004 (0.036)	0.964	0.992
Distance to night light (meters)	-0.000*** (0.000)	-0.000 (0.000)	0.105	0.000
Live in Tuma region (= 1)	-0.135*** (0.028)	-0.123*** (0.032)	0.764	0.000
Live in Madriz region (= 1)	0.128*** (0.025)	0.037 (0.028)	0.019	0.000
<i>Household Characteristics: Economic activities & Assets</i>				
Household head main occupation is ag. (= 1)	0.055 (0.057)	0.063 (0.047)	0.914	0.270
Size of landholdings ('000 sq meters)	-0.001 (0.002)	0.000 (0.001)	0.550	0.823
Number of parcels of land	0.093** (0.038)	0.030 (0.020)	0.149	0.022
Log predicted expenditures (pc)	0.013 (0.060)	-0.121* (0.066)	0.141	0.195
Wealth index - housing characteristics	-0.007 (0.015)	-0.024* (0.014)	0.399	0.205
Wealth index - productive assets	0.034* (0.017)	0.034** (0.015)	0.981	0.018
Wealth index - other assets	-0.038*** (0.013)	-0.011 (0.014)	0.173	0.020

Table B1: Correlates of Attrition. Complete Tracking Phase. (Continue)

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

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

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.²⁷ We hypothesize that a lot of the non-

²⁷We 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.

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.²⁸ 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 indigeneous 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 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.

²⁸The 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.

Table B2: Correlates of Compliance

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

Table B2: Correlates of Compliance. (Continue)

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

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

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 households, being the son of the hh head, and the number of children of the hh 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 group 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 seems very likely.

Table B3: Correlates of Attrition. Regular Tracking Phase.

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

Table B3: Correlates of Attrition. Regular Tracking Phase. (Continue)

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

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

Table B4: Correlates of Attrition. Intense Tracking Phase.

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

Table B4: Correlates of Attrition. Intense Tracking Phase. (Continue)

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

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

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.²⁹

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.³⁰ 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

²⁹Tracking results and balance for the cohort used in this paper is discussed in section 3.4

³⁰Figures 3 and 4 in Appendix C provides information of the location of all permanent migrants who were found and successfully interviewed.

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 communities. 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, 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 3: Location of Permanent Migrants found during the Regular Tracking Phase.

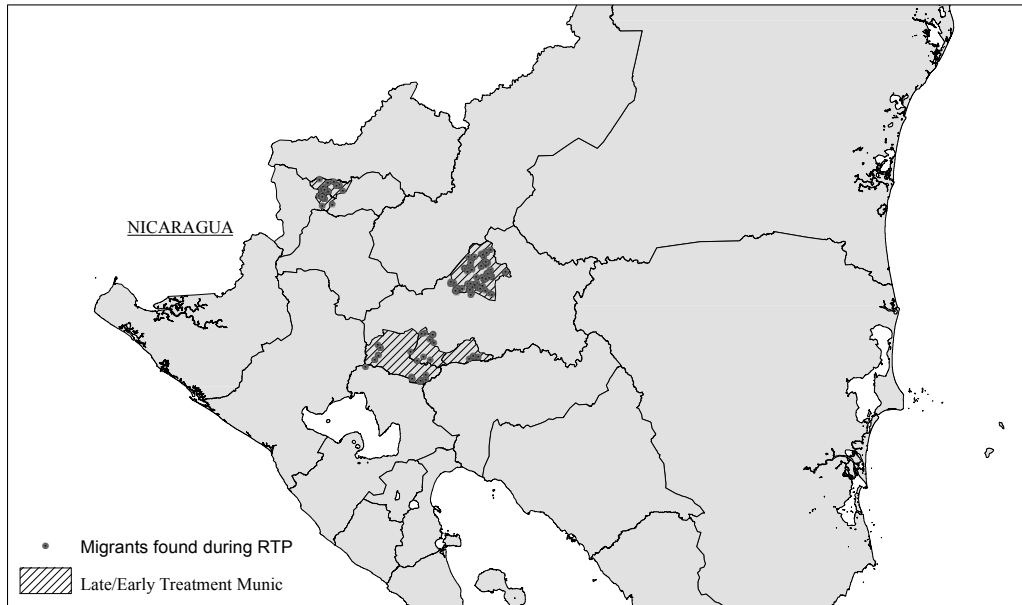
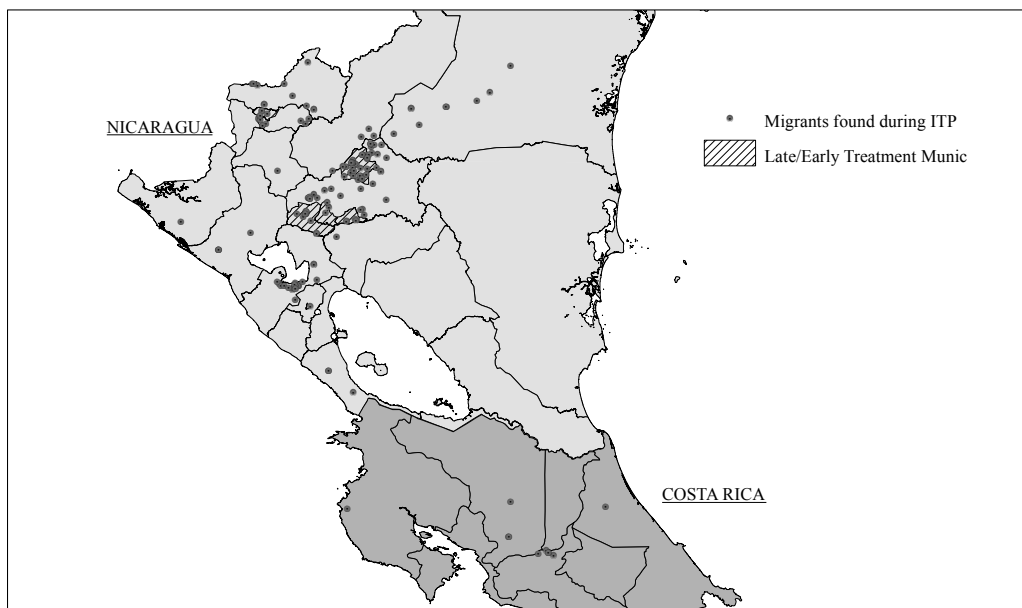


Figure 4: Location of Permanent Migrants found during the Intense Tracking Phase.



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 weight 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				Intense 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** (0.29)	0.427** (0.17)	0.282* (0.16)	0.248 (0.16)	0.865** (0.33)	0.613*** (0.20)	0.412** (0.20)	0.302 (0.18)	-0.745 (0.58)	-0.481 (0.37)	-0.527 (0.43)	-0.091 (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
<i>Comparing coefficients at different stages during the tracking process: χ^2</i>												
P-val: $T^{CTP} - T^{TP} = 0$					0.0033	0.0361	0.1293	0.5134	0.0241	0.0149	0.0581	0.3875
P-val: $T^{RTP} - T^{TP} = 0$									0.0231	0.0143	0.0580	0.3943
Off-Farm Employment												
Early-Treatment	0.059* (0.03)	0.061* (0.03)	0.056* (0.03)	0.053** (0.02)	0.087** (0.03)	0.088** (0.03)	0.087** (0.03)	0.083*** (0.03)	-0.071 (0.05)	-0.097* (0.06)	-0.147** (0.07)	-0.197*** (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
<i>Comparing coefficients at different stages during the tracking process: χ^2</i>												
P-val: $T^{CTP} - T^{TP} = 0$					0.0041	0.0076	0.0051	0.0070	0.0093	0.0030	0.00151	0.000
P-val: $T^{RTP} - T^{TP} = 0$									0.0073	0.0027	0.0012	0.000
Obs.	1006	1006	1006	1006	827	827	827	827	179	179	179	179

Standard errors are clustered at locality level. Robust s.e in brackets.*** p<0.01, ** p<0.05, * p<0.1

Notes: 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. We replace all the categorical and continuous covariates with a set of binary variables indicating whether individual is above the sample median for each of these variables. After complete tracking the off-balance controls include baseline controls for whether the individual was working, for the number of individuals with family ties in village, for the village population size and for a productive asset index (2nd principal component of household wealth). After regular tracking the off-balance controls include as well controls for whether the mother had no education, for whether the mother had at least three years of education, for whether the individual is son of the household head, for the number of children of the household head and for whether the head is female. The regression on the sample targeted during ITP the off-balance controls include baseline controls for the number of individuals with family ties in village, for the village population size and for 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.

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 (0.32)	0.363* (0.18)	0.243 (0.17)	0.232 (0.16)	0.503 (0.30)	0.392** (0.19)	0.261 (0.18)	0.229 (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** (0.32)	0.655*** (0.21)	0.501** (0.20)	0.361** (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** (0.03)	0.061* (0.03)	0.054* (0.03)	0.051** (0.02)	0.060** (0.03)	0.063** (0.03)	0.058* (0.03)	0.054** (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** (0.03)	0.086** (0.04)	0.084** (0.03)	0.081*** (0.03)
R-squared					0.03	0.05	0.07	0.09
Obs.					827	827	827	827

Standard errors are clustered at locality level. Robust s.e in parenthesis.*** p<0.01, ** p<0.05, * p<0.1

Notes: 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. We replace all the categorical and continuous covariates with a set of binary variables indicating whether individual is above the sample median for each of these variables. After complete tracking the off-balance controls include baseline controls for whether the individual was working, for the number of individuals with family ties in village, for the village population size and for a productive asset index (2nd principal component of household wealth). After regular tracking the off-balance controls include as well controls for whether the mother had no education, for whether the mother had at least three years of education, for whether the individual is son of the household head, for the number of children of the household head and for whether the head is female. The regression on the sample targeted during ITP the off-balance controls include baseline controls for the number of individuals with family ties in village, for the village population size and for 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.

F Appendix: Literature Review

F.1 Attrition Rates in RCTs

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 2015. 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.
- The analysis uses at least two rounds of data.

The final sample includes 62 articles (see below).

Table F1 presents descriptive statistics on attrition rates. We divide studies into three categories depending on the unit of analysis: children under 18 years old (21 studies), households (19 studies) and adults (21 studies). Three studies report household and individual attrition rates and one study reports attrition in 2 countries. For three papers we could not find information on the attrition rate. Table F1 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 F2 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 20 percent or higher.³¹

Most studies using panel data to evaluate a RCT report attrition incidence by treatment arms: still 15 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

³¹For readability, Table F1 also excludes two studies with annual attrition rates above 100 percent. Both studies collected follow-up data 4-6 weeks after the baseline and have attrition rates around 20 percent.

treatment arms (including between treatment and control). Seventy five 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. In 22 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, 19 percent of the studies report whether baseline characteristics for the subsample of respondents are balanced after attrition, and 37 percent consider whether selection into attrition is driven differently by baseline characteristics among treatment groups (in some cases the authors include treatment interactions). In 18 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 11 studies that report non-random attrition 7 do not apply any method to correct for non-random attrition. The other four apply IPW or non-parametric bounds. In another four 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. Thus it is becoming more common in the literature to report bounds for the range of treatment effects. Almost 16 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 forty 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. Forty percent of the studies report worst-case scenario bounds ([Horowitz and Manski, 2000](#)) and less extreme bounds as in [Kling, Liebman and Katz \(2007\)](#) which depending on the assumptions made could end being very narrow. Finally, twenty percent report bounds following the three methodologies.

Table F1: Annual Attrition Rates in the Sample Studies

	Children under 18 years old	Household	Adults above 18 years old
Average	11.64	8.6	19.39
Standard Deviation	12.97	9.11	20.49
Minimum	0.6	0.8	1.3
Median	9.3	4.27	12.3
Maximum	60	30	75
Number of studies	21	19	21

Notes: We exclude two studies for which follow-up data was collected 4-6 weeks after baseline

Table F2: Accumulated Attrition Rates in the Sample Studies

	Children under 18 years old	Household	Adults above 18 years old
Average	14.79	8.91	17.9
Standard Deviation	9.68	7.55	12.01
Minimum	1.5	1	1.3
Median	14	7.05	19
Maximum	35	25	54
Number of studies	21	20	22

Table F3: Reporting and Dealing with Attrition in the Sample Studies (62 studies)

	Report Attrition by Treatment Arms (mean comparison)	Find Balanced Attrition Rates (t-test)	Do not Proceed any Further	Study		
				Determinants of Attrition	Balanced Randomization After Attrition	Heterogeneous Treatment Effects
Number of studies	53	40	12	23	12	20
% of total studies	85.4	64.5	19.3	37.1	19.3	32.2
% of studies col. 1		75.5	22.6			

Table F4: Methods to Correct for Attrition in the Sample Studies (62 studies)

	IPW	Non-parametric Bounds	IPW+	Proxy Information	Proxy+Heckman Correction model	None
Number of studies	6	8	2	1	1	44
Selection on observables:						
<i>Random Attrition</i>	4	2	0	0	0	10
<i>Non-random Attrition</i>	1	1	2	0	0	7
<i>No reported</i>	1	5	0	1	1	27
% of total studies	9.6	12.9	3.2	1.6	1.6	70.9

List of papers in the review

- 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.
- 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.
- 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.
- 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.
- 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.
- 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.
- 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.
- 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.
- 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.

- 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.
- 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.
- 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.
- Chinkhumba, Jobiba, Susan Godlonton, and Rebecca L Thornton.** 2016. “The Demand for Medical Male Circumcision.” *American Economic Journal: Applied Economics*, 6(2): 152–177.
- 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, 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.
- 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.
- 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, Rema Hanna, and Stephen P Rya.** 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.
- 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.
- 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, 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.
- Glewwe, Paul, Nauman Ilias, and Michael Kremer.** 2010. “Teacher Incentives.” *American Economic Journal: Applied Economics*, 2(3): 205–227.
- 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.
- 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.
- 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.
- 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.
- 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.
- Kremer, Michael, Edward Miguel, and Rebecca Thornton.** 2009. “Incentives to Learn.” *The Review of Economics and Statistics*, 91(3): 437–456.
- 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.
- 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.
- 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.

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

F.2 Attrition Rates in Longitudinal Surveys

Table F5: 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 F6: 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 Weerd, 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 Weerd, 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

tion: 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)*.