HAL archives-ouvertes.fr

Education, Income and Mobility: Experimental Impacts of Childhood Exposure to Progresa after 20 Years

Maria Caridad Araujo, Karen Macours

► To cite this version:

Maria Caridad Araujo, Karen Macours. Education, Income and Mobility: Experimental Impacts of Childhood Exposure to Progresa after 20 Years. 2021. halshs-03364972

HAL Id: halshs-03364972
https://halshs.archives-ouvertes.fr/halshs-03364972

Preprint submitted on 5 Oct 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

PARIS SCHOOL OF ECONOMICS
ECOLE D'ECONOMIE DE PARIS

WORKING PAPER N° 2021 – 57

Education, Income and Mobility: Experimental Impacts of Childhood Exposure to Progresa after 20 Years

M. Caridad Araujo
Karen Macours

JEL Codes:
Keywords:

AGENCE NATIONALE DE LA RECHERCHE
ANR
INVESTIR L'AVENIR
Funded by a French government subsidy managed by the ANR under the framework of the Investissements d'avenir programme reference ANR-17-EURE-001

Education, Income and Mobility:

Experimental Impacts of Childhood Exposure to Progresa after 20 Years

M. Caridad Araujo and Karen Macours[1]

Version: October 4, 2021

Abstract

In 1997, the Mexican government designed the conditional cash transfer program Progresa, which became the worldwide model of a new approach to social programs, simultaneously targeting human capital accumulation and poverty reduction. A large literature has documented the short and medium-term impacts of the Mexican program and its successors in other countries. Using Progresa's experimental evaluation design originally rolled out in 1997-2000, and a tracking survey conducted 20 years later, this paper studies the differential long-term impacts of exposure to Progresa. We focus on two cohorts of children: i) those that during the period of differential exposure were in-utero or in the first years of life, and ii) those who during the period of differential exposure were transitioning from primary to secondary school. Results for the early childhood cohort, 18–20-year-old at endline, shows that differential exposure to Progresa during the early years led to positive impacts on educational attainment and labor income expectations. This constitutes unique long-term evidence on the returns of an at-scale intervention on investments in human capital during the first 1000 days of life. Results for the school cohort - in their early 30s at endline - show that the short-term impacts of differential exposure to Progresa on schooling were sustained in the long-run and manifested themselves in larger labor incomes, more geographical mobility including through international migration, and later family formation.

[1] Araujo: Inter-American Development Bank (mcaraujo@iadb.org); Macours: Paris School of Economics and INRAE (karen.macours@psemail.eu). This paper extends the 2018 (Spanish-language) working paper titled "Efectos de largo plazo de la exposición diferenciada a Progresa". We gratefully acknowledge the inputs and support of Prospera's evaluation team and in particular Josué Vargas, Rogelio Grados, Raúl Pérez, Martha Cuevas, and Susana Torres throughout the project, as well as support and comments from Pablo Ibarraran, Santiago Levy, Ferdinando Regalia and Norbert Schady, and seminar participants at Columbia, Essex, Gottingen, Georgetown, IADB, IFS, Javeriana, Leuven, Paris 8, PSE, Stanford, Tilburg, TSE, UC Davis, Universidad de Los Andes, Vanderbilt, and the World Bank. Juan Nicolás Herrera, Marta Dormal, María del Carmen Hernández, and Carolina Rivas provided excellent research assistance at different stages of the project. We also thank José Marín Casiano and the Berumen Asociados field team for all data collection and tracking efforts which were financed by the Inter-American Development Bank. Data collection was approved by PSE-JPAL Europe IRB, nr 2015-005. The views expressed in this paper are those of the authors and do not necessarily reflect those of the Inter-American Development Bank, its Executive Directors, or the countries they represent, or any of its affiliated organizations. All errors and omissions are our own.

1. Introduction

Policies aimed at increasing human capital, whether in early childhood or during schooling years, remain an important global priority, explicitly recognized by the United Nations Sustainable Development Goals. Over the last 20-plus years, there has been remarkable progress in a number of human capital indicators across the world. There has also been a remarkable level of policy innovation, often based on micro-economic insights regarding households' decisions on human capital investments, paired with efforts to rigorously evaluate new policy designs and approaches. These advances bring hope that those who benefitted from such interventions as children, may start their adult lives better prepared than previous generations. At the same time, large concerns remain both because negative shocks later in childhood or adolescence (including global pandemics) can erase early gains and given mounting evidence that more educational attainment often does not result in gains in learning (World Bank, 2018).

Direct empirical evidence on whether the human capital gains obtained as a result of specific policy interventions in low- or middle-income countries persist in the long run and lead to better adult life outcomes is limited. There are a number of notable positive results of interventions targeting school-aged children (Duflo, Dupas, and Kremer 2015, 2021; Hamory et al, 2020; Bettinger et al. 2016), though not all studies observe people once they have fully transitioned into adulthood. While the medical literature highlights the importance of the first 1,000 days of a child's life for human brain development and physical growth (Grantham-McGregor et al, 2007), there is little experimental impact evidence of interventions in the early years on outcomes in adulthood, with the most prominent long-term evidence coming from relatively small-scale stimulation (Gertler et al. 2014; 2021) or nutrition interventions (Hoddinott et al. 2008).[2]

Questions on whether hypothesized long-term returns materialize are particularly relevant for human capital gains obtained as a result of anti-poverty programs. The populations targeted by these programs are likely to face adverse conditions after childhood, possibly offsetting some of the early gains. Moreover, in many countries, gains in human capital among the poor often occur in an environment where many others -especially the non-poor- also obtain higher levels of human capital. Thus, increased schooling amongst the poor may not automatically translate in labor market advantages, in particular if learning is limited. Schooling can, however, affect income and long-term welfare through other channels. These questions are particularly relevant for conditional cash transfer (CCT) programs, a policy innovation that was scaled globally. More than 20 years after the first CCTs started, enough time has passed for the first generations of beneficiaries to have fully reached adulthood.

---

[2] There is much more extensive evidence of the importance of nutrition and health shocks during the early years of life on later outcomes leveraging natural shocks or quasi-experimental policy changes (Almond et al, 2018).

This paper provides 20-year experimental evidence on the long-term impacts of human capital investments in the first years of life, and during school age, of the first large-scale and rigorously evaluated CCT, the Mexican program Progresa. The program´s multisectoral approach intervening in health, education and nutrition simultaneously (Skoufias, 2005) allowed for possible dynamic complementarities of investment in the early years with further human capital investments, emphasized by Heckman (2006). We analyze whether short-term impacts of differential exposure to the CCT during periods often considered critical for human capital formation (early childhood and transition from primary to secondary school) translated into better demographic, human capital, and labor market outcomes. We use original data obtained through tracking and interviewing close to seven thousand individuals from the initial evaluation sample.

Progresa had a rigorous evaluation embedded in its design. Informed by the results of the initial experimental evaluation, the program was scaled up nationally starting in 2000 and served as model for many similar programs in Latin America (covering 25 percent of the population in region) and beyond, to more than 80 countries worldwide. The initial experiment took place in 506 villages, of which about 2/3 (320) were randomly selected as treatment villages where transfers started in May 1998, while the other 1/3 (186) served as control until November 1999, when they began receiving transfers. The program short- and medium-term impacts have been extensively studied. Considering the impacts of the health and nutrition components of the program, short-term evaluations showed positive impacts of Progresa on health and nutrition investment, and on children's health outcomes, including height gains and a reduction of the prevalence of anemia during the early years of life (Gertler, 2004; Rivera et al., 2004; Behrman and Hoddinott 2005). Regarding the educational components, Schultz (2004) established that the short-term impact on the likelihood of staying in school was particularly large during the transition from primary to secondary school. Evidence from the medium-term evaluation showed there was still a differential effect on schooling six years later, but not on learning (Behrman, Parker, Todd 2009a, 2011).

This paper focuses on two cohorts of children who, during the period of experimental differential exposure to Progresa, were of ages considered particularly important for human capital formation: either very early in life or in the transition from primary to secondary school. Throughout the paper, we refer to them as the *early childhood cohort* and the *school* cohort, respectively. The *early childhood cohort* are children born between 1997 and 1999. A child born during this period in the early treatment group (T1 hereafter) benefited from Progresa during the first 1000 days of life (including in utero for those born late in 1998 and in 1999). In turn, a child born during that period in the experimental control group (T2 hereafter) only started receiving the program after her first year(s) of life. We observed this cohort when they are 18 to 20 years old in 2017, and other than one group starting earlier than the other, they both would have been exposed to the program throughout their childhood. In 2017, they were transitioning into adulthood: some were still studying while others had completed their education and were entering the labor force and/or starting a family. To estimate long-term returns beyond education, we measured income expectations (Attanasio and Kaufmann, 2014), circumventing the challenge of inferring impacts during transition into adulthood, when income is still censored by construction.

The *school cohort* consists of children randomly exposed to Progresa when they were transitioning from primary to secondary school. These are children who were in 6th grade of primary school in 1998. Those living in T1 villages would have received Progresa during their transition to secondary school, while their peers living in T2 villages would have received the program 1.5 years later when this transition normally had already happened. They would be more likely to have dropped out instead of transitioning to secondary school. Individuals in this cohort are between 29-35 years old in 2017 and would have completed their education, transitioned into the labor market and, typically, formed their own independent households.

We intensively tracked children in these two cohorts from the original evaluation sample of eligible households, collecting information on 93 percent of them - with no difference in tracking rates between T1 and T2.[3] The majority of individuals had migrated out of their locality of origin and had moved to a broad range of destinations in Mexico and the US. Such geographic mobility is a key outcome of interest to understand long-term returns, but also poses challenges to many long-term studies that the tracking helped avoid.[4] Descriptively, the data also shows a remarkable level of intergenerational educational mobility.

Across the two cohorts studied, exposure to the program at critical ages in childhood leads to increased schooling attainment after 20 years and also manifests in other adulthood outcomes. Results show positive impacts on schooling of a modest order of magnitude of 3 to 4 percent, equivalent to 0.3 to 0.4 additional years. Earlier exposure has positive impacts on the likelihood of completion of lower secondary school (9th grade) of 6 to 12 percent, and for the early childhood cohort also on the likelihood of completing upper secondary (12th grade) and having university studies, with increases of 16-67 percent. For both cohorts, differential exposure translates in higher geographic mobility 20 years later, with less individuals from T1 villages still living in their village of origin (compared to T2), and more likely to have migrated out of the municipality. For the early childhood cohort, many of whom are still studying, we show an 8 percent increase in income expectations. For individuals from the school cohort - in their early thirties at endline - earlier exposure leads to 16 percent increase in labor incomes, more migration to the United States and US dollar income, as well as a delay in parenthood by approximately half a year. While there is no strong evidence of differential effects by gender, point estimates for educational outcomes and income(expectations) are larger for women in both cohorts.

[3] Fifty percent of the individuals in the tracking sample were interviewed in person and 35 percent were interviewed by phone. Information collected on the remaining 8 percent was through proxy informants. People found through these different tracking methods exhibit significantly different characteristics, confirming the value of using a combination of search strategies during fieldwork. Tests for the reliability of information obtained through the different methods were built into the data collection and drawn on in the analysis.

[4] Both with other survey data and with administrative data sources the ability to fully account for international migration in particular is often limited.

This paper builds on the large body of evidence of short-term impacts on Progresa (reviewed by Parker and Todd, 2017) and on literature on the medium-term (5 to 13 year) impacts of CCTs more broadly (reviewed by Molina-Millan et al., 2019). Evidence on impacts of exposure in early childhood is limited, with some papers documenting fade-out, while others find cognitive gains or gains in early schooling. More evidence exists on school-age exposure, showing almost everywhere gains in schooling but mixed and inconclusive evidence on labor market outcomes, earnings, and migration (Araujo, Bosch, Schady, 2018; Baird, McIntosh, Ozler, 2018; Barham, Macours, Maluccio, 2013, 2018a,b; Cahyadi et al, 2020; Molina-Millan et al, 2019,2020). Most of this research focuses on outcomes at ages when individuals were still transitioning from schooling to work and had not formed their own households. Our paper underscores the value of observing 20-year outcomes for individuals in the school cohort when they were around 30 years old and had fully transitioned into adulthood. While the experimental design does not allow to estimate the absolute program impacts, we show that differences in exposure led to long-term human capital gains, geographic mobility, and income, as such uncovering mechanisms through which the CCT-induced human capital investments affect long-term trajectories.

Our analysis for the school cohort most directly complements Parker and Vogl's (2019) non-experimental analysis of the 13-year absolute impacts of the program. Using a sample of 10 percent of all households in the 2010 national population census and an identification strategy based on the spatial and time variation in Progresa's national rollout and expansion, their difference-in-difference estimates show effects of 1.5 more years of schooling for men and women of primary-school age at the start of the program. Results further show increased labor force participation by 30 to 40 percent and labor incomes by 50 percent, with labor market impacts found to be larger for women than for men.

The contribution of our paper is three-fold. We provide the first at-scale experimental impact evidence of a national level program increasing investments starting in the first years of life, on educational attainment 20 years later when former beneficiaries enter adulthood, and document how it translated into income expectations and geographic mobility. While many interventions in early childhood are motivated by the potential of long-term returns, supporting experimental evidence has been limited to long-term follow-ups of relatively small-scale RCTs, implemented by researchers. Our results on a program implemented by the government in real-life conditions, that has served as model for similar programs across the world, constitutes a valuable addition to the evidence base. Second, we provide the first experimental evidence of a CCT program on income and on other life outcomes once former beneficiaries have fully transitioned into adulthood. Third, we demonstrate the importance of geographical mobility as both a pathway and an outcome of interest when analyzing the long-term returns to human capital investments, as well as the feasibility of tracking individuals to better understand dynamic pathways over 20 years. We highlight the possibility that schooling may directly help expand networks and increase geographic mobility. In rural settings, completing higher levels of education often implies completing schooling outside of the village of origin, and hence mechanically may increase students' mobility, and expose them to possibly better-connected networks, a key asset for future migration.

To interpret the findings, the overall context in Mexico during the two decades since the late 90s is important to consider. There were large increases in educational attainment nationally – but also large frictions in the labor market and substantial and increasing misallocation, limiting the aggregate returns to the gains in human capital (Levy, 2018).[5] The evidence in this paper shows that, even with such frictions, human capital investments by poor rural households led to positive returns, and suggests that geographic mobility can help explain these findings. As such we contribute to the literature on the relationship between education, geographic mobility, and migration.

The inability of entrepreneurs and workers to move to locations where they are the most productive, both within a country and internationally, is a long-recognized source of inefficiency. Networks have been shown to help overcome some of the frictions limiting geographic mobility of workers (Munshi 2020 provides a recent review), and there is specific evidence of their importance for Mexico-US migration (Winters et al, 2001; McKenzie and Rapoport, 2007; Munshi, 2011) and for migration within Mexico (Davis et al, 2002). Migration networks can also help explain why some communities are more likely to send less educated individuals to the US, while positive selection is found for others (McKenzie and Rapoport, 2010). While there is a debate on whether selection into Mexico-US migration is positively or negatively related to human capital (Chiquiar and Hanson 2005; Caponi 2011), there is little evidence on how exogenous increases in schooling affect migration patterns. Return-migration from the US to Mexico is also a phenomenon that has been less studied.[6] Our work is suggestive of its contribution to economic mobility (in line with Li 2017, who explores how length of stay abroad affects earnings at home).

This paper more broadly relates to the literature on the rural productivity gap, the spatial misallocation of labor, and selective migration in developing countries (Young, 2013; Gollin, Lagakos, and Waugh, 2014; Lagakos, 2020; Lagakos et al 2020; Hamory et al 2020), the returns to migration (Beegle, De Weerdt and Dercon, 2011; Bazzi et al, 2016; Bryan and Morten, 2019), and to interventions lifting constraints to migration (Bryan et al 2014; McKenzie et al 2010; Baseler 2021). While the relationship between returns to human capital and location is central to this broad literature, our contribution focuses on providing evidence of the relationship between an intervention that increases human capital investments in childhood and geographic mobility later in life.

[5] Levy (2018) discusses how various policies and institutions in Mexico together lead to a sub-optimal distributions of individuals across occupations and of firms across sectors or sizes, and to sub-optimal matches between firms and workers of different abilities.

[6] Work on return migration from the US to Mexico has focused mainly on the health status of return-migrants (Donato, Hamilton and Bernard-Sasges 2019; Waldman, Wang and Oh 2019; Ullman, Goldman and Massey 2011 and Wilson, Stimpson and Pagán 2014). Lindstrom (1996) also hypothesized that return migration in the 1990s was more frequent amongst individuals whose locations of origin in Mexico were poorer and stagnant and thus where savings from migration would have fewer productive returns.

This evidence on the impact of a CCT on long-term migration also relates to evidence on how the 10-year returns to a large asset transfer programs increased through migration (Banerjee, Duflo and Sharma, 2021).

Our main empirical strategy focusses on two specific cohorts, as short-term evidence and other literature point to these cohorts as being the groups for which there is the most statistical power in the differential experimental design. It leaves unanswered whether other cohorts benefitted in similar ways. To partly address that question, we collected a limited set of outcome variables for the siblings of the individuals tracked, covering age ranges between 1 and 17 years at baseline. For both the sibling sample and the main sample we also compare our results with the medium-term results, allowing to address possible concerns regarding selection of the sibling sample. We show that differential effects on education for siblings who started benefitting at an older age than the early childhood cohort (i.e., those between 1 and 8) are positive but smaller in size than those of the early childhood cohort. Differential effects for siblings between 9 and 15 that were either lagging behind the main school cohort or were not at all enrolled at baseline show that the program helped them attain primary school completion, while those further ahead or older exhibit no differential educational gains. Interestingly, among these groups we find no increased geographic mobility nor international migration, suggesting that the geographic mobility results for the main cohorts are not primarily driven by the higher cash transfer amounts received by T1 households.

The remainder of the paper is organized as follows. Section 2 reviews the program and initial experimental design, as well as the short-term evidence on human capital investments and outcomes. Section 3 discusses the sample design and tracking of the 20-year follow-up data collection and shows descriptive evidence on educational and geographic mobility**.** Section 4 introduces the empirical specification, and section 5 discusses the main results for both cohorts. Section 6 focuses on the differential impacts for the siblings of the individuals of the main sample, while section 7 concludes.

2. Program design and short-term evidence

Created in 1997, Progresa (later renamed Oportunidades and then Prospera) originally operated only in rural areas. The core elements of the program consisted of sizable transfers to mothers of eligible households, conditional on a) health check-ups and nutritional supplements to pregnant mothers and young children (0-5); b) school enrollment and attendance by children in grades 4 to 6 of primary school and in lower secondary school (grades 7 to 9); and c) attendance to information sessions on nutrition, health and education practices. Mothers received a fixed amount for nutrition, complemented with grants for school-going children, that increased with the grade-level, and were slightly higher (13 percent) for girls than for boys. The program was means tested with a village-level marginality index used for geographical targeting, and a household-level level proxy means test for household targeting based on poverty. This is reflected in the baseline outcome indicators of the targeted children, showing, for instance 44% of 1-3 year olds being stunted (Skoufias, 2005).

Skoufias (2005) and Parker and Todd (2017) provide a comprehensive review of the program, the evaluation design, and the literature documenting the various impacts of the program. Indeed, part of what drew the international development community´s attention to Progresa was that it used its gradual rollout to conduct an experimental impact evaluation. Localities were randomly assigned into a control group (N=186) and a treatment group (N=320), and eligible beneficiaries from the former started receiving the program approximately 18 months later than those from the latter. All households (over 24 thousand) in these 506 localities were followed in a panel of upto eight surveys conducted between 1997 and 2007 (i.e., the ENCASEH survey conducted in November 1997 and the ENCEL surveys conducted in October 1998, March and November 1999, March, and November 2000, 2003, and 2007).[7] After 2007, no further data was collected on this sample till 2017.

Initial take-up was very high (90-95 percent), and families continued to receive transfers for up to 20 years. Apart from the difference in starting date, the experimental variation also (mechanically) implies a difference in the total amount received by these households over the 20-year period – (Appendix A documents compliance and program exposure for the two cohorts). In 2001, educational grants were extended to upper secondary school (grade 10-12) and in 2017, to tertiary education.

The short-term impact evaluation results showed that children from the early childhood cohort in T1 villages had better health outcomes (fewer illnesses), lower prevalence of anemia and were taller than those from T2 villages. There is also evidence on earlier prenatal visits, increased investments in health and nutrient-rich food (Gertler, 2004; Behrman and Hoddinott 2005; Rivera et al. 2004; Skoufias 2005). Using variation in the national roll-out, Barham (2011) shows Progresa reduced infant mortality by 17 percent but had no impact on neonatal mortality. Fernald, Gertler and Neufeld (2009) explored the persistence of the impacts of differential experimental exposure by analyzing data collected ten years later, in 2007, in the rural evaluation localities (but not among those that migrated out). They found reduced prevalence of behavioral problems but found no sustained differential effects on nutritional, cognitive development or language outcomes.

Of the three birth-year groups in the early childhood cohort (1997, 1998, and 1999) those born in 1999 would have been conceived after the start of the program. While the short-term evaluation results showed no significant fertility effect in the first 18 months of the program (Stecklov et al., 2007), we cannot exclude

[7] In 2003 and 2007 only limited tracking outside of the original 506 localities was done, resulting in attrition rates of upto 60% for the school cohort. Kugler and Rojas (2018) combine the 1997, 2003 and 2007 rounds of the ENCEL rural survey with program administrative data collected between 2008 and 2015 to study the relationship between school-age exposure to Progresa and duration of benefits on educational and labor market outcomes. This administrative data only includes individuals still living in original baseline households and only if those households are still in the program, resulting in loss of more than 95 percent of the original sample.

possible heterogeneous fertility effects. We therefore also present all analyses only including the 1997 and 1998 cohorts. We test for differences between birth-cohorts, to explore if exposure at different moments within the first 1000-day window made a difference.

For the school cohort, the short-term impact evaluations showed that Progresa positively affected the likelihood of transitioning from primary into secondary school. In the short term, children assigned in treatment localities attained 0.66 years more years of schooling than those in the control ones (Schultz, 2004). A follow-up study of these same children showed that six years later, the differential exposure to Progresa translated in modest differential impacts on schooling. Children at high-risk of dropout, i.e., those transitioning from primary to secondary school during the period of differential exposure, attained 0.5 extra years of school. However, there were no significant differences between the early and late treatment group on grade progression for younger children, nor on writing, math or reading tests for the older cohorts (Behrman, Parker and Todd, 2009a,b).[8]

The encouraging short-term evidence from Mexico led to a rapid adoption and scaling-up of CCTs in several countries in the region. By 2013, 137 million persons in 17 countries in Latin America and the Caribbean were receiving cash transfers (Ibarraran et al. 2017). Key elements of the design and implementation of Mexico´s Progresa, such as the use of a proxy means targeting formula, the structure of schooling and health conditionalities, the variable nature of the transfers, amongst others, influenced the design of programs elsewhere.

3. Sample Design, Tracking and Mobility

Sampling

To identify the children in the two cohorts of interest that would be tracked and interviewed about their outcomes 20 years later – henceforth referred to as the tracking sample - we started with the original data from the impact evaluation, that comprised over 24,000 (eligible and non-eligible) families in 506 villages. These villages were randomly selected from 4,546 eligible villages in seven states in Central Mexico (Figure 1), stratified into five geographical regions and four village population-size categories, with equal shares of treatment and control villages drawn in each (Progresa, 1997; Bobba and Gignoux, 2019).[9]

The ten-year follow-up survey in 2007, which aimed to track all 24,000 households, had an attrition rate of 60 percent for some cohorts, including the school cohort in this paper. Ten more years later, with limited contact

[8] Little is known to date, on the migration impacts for the children who benefitted from the human capital components of the program during childhood. Two short-term impact studies do however study the impact of the cash inflows on migration for older cohorts (parents of the targeted children), showing that migration to the US was higher in T1 than in T2 in the first year of the program, but that this had reversed after 18 months, while there were no significant short-term impacts on domestic migration (Stecklov et al, 2005; Angelucci, 2015).

[9] Eligible villages in those seven states typically were located away from the Coast and large cities, many of them are in mountainous areas and have a relatively large share of indigenous population.

information available other than the village of origin, an estimated 20 percent of original beneficiary households having entirely moved out of those villages, and a larger share of individual-level migration expected for those who benefitted from the program and were of particularly mobile ages in 2017, we anticipated attrition to be an even bigger challenge and purposely organized the data collection to limit it. Concerns about attrition were compounded with concerns about statistical power. The initial experimental design leads to a comparison of 18.5 years of program exposure to 20 years of exposure, raising questions on whether this relatively small difference provides enough statistical power. To address these two challenges, we decided to focus the tracking sample on two narrowly defined cohorts for whom the differential timing of exposure occurred during moments considered critical for human capital accumulation in childhood. [10] We then tracked intensively all individuals in those cohorts. [11]

To identify children in the early childhood cohort, we used the panel of the 1997 ENCASEH and the first 6 waves of the ENCEL surveys (1998 to 2000) and selected all children born between 1997 and 1999 in households eligible to receive the program. Identification of the children in the school cohort was done based on data from the 1997 ENCASEH, from which we selected all individuals in eligible households that were in primary school and had completed 5th grade by the time of November 1997 ENCASEH. Transfers started in May 1998, so they would have received the first transfers just before making the decision to enroll in lower secondary in July 1998. Appendix B provides further details on sampling.[12]

Migration tensions with the US and domestic security threats imposed additional constraints for the data collection. The tracking survey took place in 2017, when there was a climate of insecurity and ongoing criminal (drug) gang violence in many parts of Mexico. Moreover, there was heightened uncertainty and apprehension given the expected crack down on Mexican migrants after the election of Donald Trump in the US. Of the 506 villages in the sample, there were 47 villages where the survey firm and Prospera staff agreed it would be impossible to conduct fieldwork due to insecurity. Asking for contact information in settings with active drug gangs would have implied risks for both the survey teams and respondents. While these 47 villages were somewhat poorer than the rest of the sample at baseline, there was approximately the same proportion of T1 (30 villages) and T2 (17 villages) as in the full sample. As checks with baseline variables confirmed that excluding

[10] Among the households initially classified as non-eligible in T1 and T2, many became eligible for the program soon after its start. Several papers have also shown substantial spillover effects on these non-eligible households. For both reasons, a possible alternative approach to estimate long-term impacts using a RDD around the household-level eligibility threshold, would be hard to interpret.

[11] The tracking survey was purposely organized in 2017 by Prospera's evaluation team with support of the IADB to generate new evidence on possible long-term impacts of the program, as it reached its 20-year anniversary. Results were presented to policymakers in Mexico and made available in an earlier (Spanish-language) version of this paper. The tracking survey for the two cohorts was followed by a separate effort by Prospera in 2017-2018, that aimed to re-interview all the original households in a subset of villages of the ENCEL panel, as well as a subset of their split-off households.

[12] The academic year in Mexico runs from late August to early July.

these villages would not affect balance, no interview efforts were made there.[13] Most (39) of these villages were located in Guerrero and Michoacán (the two states were each considered separate regions in the 1997 sampling frame), while the remaining were in Veracruz. To test sensitivity of the findings to the exclusion of these villages, we also estimate all effects on the subsample of strata in which no villages were excluded (13 of the 20 strata formed by combining region and population-size categories). This is costly in terms of statistical power, but a useful check on internal validity. In addition, we show results that overweight the observations in surveyed villages in the same strata to restore representativity of the original sample.

We re-estimated the 2003 medium-term impacts on education on the sample we tracked in 2017. Table C1 shows results closely mirror (both in terms of significance and point estimates) those on the full sample of 506 villages, documented also in earlier work. In particular, for the school cohort, the differential timing of exposure led to a significant increase in the probability of finishing lower secondary school but no impacts on other levels of education. Appendix C provides further details.

Tracking

Overall, we attempted to track 6,750 unique individuals from the Progresa impact evaluation sample from two cohorts, 4,461 from the early childhood cohort and 2,289 from the school cohort. These individuals were members of 5,468 households, in 456 villages.[14] The survey took place in two phases (May-August 2017, and Sept-Dec 2017) and combined multiple methods to obtain information, starting with visits to the baseline villages and households for face-to-face interviews, to searches in all destinations in Mexico (in the second phase) for face-to-face interviews, as well as interviews by phone to all reachable migrants, including those living in the US. Different strategies were combined to maximize response rates. We leveraged the social capital of the program and, in particular, of its volunteer leaders in the village of origin, networks of migrants in destinations, parents of migrant children, assistance of parents to help gain confidence and contact information (and coordinated follow-up on "hot links" to maximize success), contacts within extended family networks, multiple visits, and a video message explaining the importance of the survey, shown in person and shared through WhatsApp. Individuals (and/or their family in the village of origin) not tracked in the first phase received a personalized invitation letter with information about the objectives of the survey prior to second phase tracking. Appendix D provides details on the survey protocols.

The survey was successful in collecting information about most of the individuals in the sample. Surveys were conducted for 93 percent of the individuals in the two cohorts (93 percent from the early childhood cohort and

[13] While exposure to the CCT program is likely to have affected village dynamics (including possibly the conflict environment), by 2017, both T1 and T2 villages had been exposed to the program for a very long time (18.5 and 20 years), which can help explain why there was no difference in the probability of violence.

[14] Apart from the 47 villages mentioned earlier, there are 3 villages where there were no children in the cohorts of interest.

94 percent from the school cohort). Of the full sample, 50 percent was interviewed face-to-face (55 percent for the early childhood cohort and 40 percent for the school cohort) and phone interviews allowed us to reach an additional 35 percent (31 percent for the early childhood cohort and 43 percent for the school cohort). The information collected on the remaining 8 percent was obtained from proxy informants (i.e., a member of the baseline household, typically the mother).[15] Among those for whom no survey could be conducted, 2.3 percent had died, 1.3 percent refused or was unable to answer, and 3.2 percent were no longer living in their villages of origin and no proxy or contact information could be obtained.[16] Importantly, there were no differences in the success of tracking across individuals from villages randomly assigned to T1 or T2, and this holds when excluding information obtained through proxy reports, or by phone (Table 1).[17]

Table D1 shows that the total sample, the sample of individuals with information (collected in-person, by phone or through proxy informants), and the sample of individuals interviewed directly (in person or by phone) are balanced on baseline characteristics for both cohorts. The balance tables also depict the conditions in which the households in our sample lived in 1997 which resemble those of many poor rural households across the developing world today. About 20-25 percent lived in a house with cardboard roof, only 25 percent had access to running water, only 15-25 percent of household heads had completed primary occupation, about 60 percent of household heads was working as agricultural laborers, and 44 percent was indigenous. While Mexico is a middle income county, the children in this sample grew up in conditions that were worse than in many other places in the country.

The 2017 Survey Instrument

The survey used a short questionnaire, designed to be administered either in face-to-face or in phone interviews. To accommodate the latter, the survey length was kept to approximately 20 minutes. It focused on key variables of interest: education, geographic mobility, occupation, and demographic information (fertility, marriage status). Test-retest statistics confirm that household informants could accurately report on these outcomes. Appendix D provides more details on the data collection.

The questionnaire was largely identical for the two cohorts. When the individuals were themselves interviewed, we also inquired about their income, more detailed information regarding their occupation, recall information

[15] Appendix D shows the reliability of the information obtained through phone and proxy surveys.
[16] These tracking rates over 20 years are on par with recent longitudinal survey efforts tracking individuals over (mostly) shorter periods, including those using intensive tracking on a random subset of the "hard to find" subsample and reweighting resulting in effective tracking rates of 84 percent after both 10 and 20 years (Baird at al., 2016, Hicks et al, 2020); 91 percent after 7 years (Duflo, Dupas and Kremer (2015) and 87 percent after 9 years (Blattman et al, 2020); Duflo, Dupas and Kremer (2021) obtained tracking rates of about 95 percent (96 percent) after 9(11) years by maintaining phone contact with respondents over the entire study period.
[17] This balance can be explained by the intensive tracking, leaving only a small share not found and the fact that households and villages in T1 and T2 all had had a long exposure to the program by 2017.

on program participation, migration history, expectations on education and income (for the early childhood cohort), and asset ownership (for the school cohort).[18] Keeping the survey short meant we could not collect detailed information on mechanisms or other welfare indicators, such as expenditure or health outcomes. Nor could we administer cognitive or achievement tests. As a proxy for health investments, we asked about smoking and drinking behavior, acknowledging this data can suffer from social desirability bias.

Geographic and educational mobility

Figure 1 and Table 2 demonstrate the geographical mobility of the individuals in the tracking sample. By 2017, they had migrated to 30 of Mexico's 32 states, to 28 US states and (one person) to Canada. Of the school cohort, only one third still lived in their village of origin, 40 percent had moved out of their home state, and 12 percent was in the US. Migration destinations in Mexico vary not just geographically, but also in the type of destinations, with 18 percent in the three largest metropolitan areas (Mexico City, Monterrey, and Guadalajara), but also about 30 percent in other urban and semi-urban municipalities each. About half of the early childhood cohort still lived in the village of origin and many at 18-20 years old would still be living with their parents. Even so, 40 percent had moved out of the municipality, and 25 percent had moved out of their home state. In striking contrast with the school cohort, migration to the US was minimal. Instead, they migrated within Mexico, mostly to urban municipalities outside of the metropolitan areas and the semi-urban municipalities. Finally, a relatively small share of respondents had moved before adulthood: 6 percent reported studying in a state different from the one where they were born. Table 2 shows that the destinations among those for whom tracking was successful and the full sample for whom some information on destination was available are very similar, confirming that the intensive tracking helped avoiding selection based on destination. Despite the geographic mobility of the individuals in this sample, the success of the tracking survey reflects that most of them kept links -immediate family, extended family, or friends- in their village of origin.

We complement this descriptive analysis of geographic mobility by illustrating the intergenerational educational mobility. Table 3 compares the 2017 educational outcomes of the two cohorts in the tracking sample with those of the household head at baseline, typically the father of the beneficiary. The first column shows the 1997 education levels of all household heads in the sample. They had, on average, three years of education, with 24 percent having completed only primary school, and almost nobody having finished any higher level. When limiting the sample of household heads to those approximately 30 years old (to facilitate comparisons with the school cohort in 2017), these numbers are slightly higher (4 years of education and 41 percent with complete primary). Twenty years later, the educational distribution looks very different: the school cohort has, on average, nine years of education (three times more than their fathers). Virtually everybody finished primary school, two-

[18] Information on ownership of ten assets is aggregated in an asset index using principal components analysis.

thirds completed lower secondary school and 6 percent started tertiary education. Further increases are observed amongst the early childhood cohort, with 35 percent of them having finished upper secondary school. Moreover, 23 percent of that cohort were still studying when the survey was conducted, so their final educational outcomes will be larger than those observed in 2017.

The descriptive analysis in Table 3 suggests a remarkable intergenerational increase in educational levels. This is not necessarily the result of exposure to the CCT. Over the same period, many others in Mexico were also getting more education. For individuals of the ages of the school cohort, average schooling nationally was 10.4 years in 2018 (compared to 8.8 years in our sample), with national averages for those born in 1997 to 1999 (the early childhood cohort) was 10.3 years (compared to 9.5 in our sample). [19] Therefore, it is not obvious that the individuals of our sample, who are coming from very marginal communities in the country, would be competitive in the labor market, even if the program helped to improve their education levels. This is particularly so given concerns with the quality of education in such settings (resulting possibly in limited learning), and the well-known frictions in the Mexican labor market (Levy, 2018). The experimental variation obtained from the difference in program exposure between T1 and T2 can help answer these questions.

4. Empirical specification

We estimate the following individual-level model, separately for each cohort:

$$Y_{il} = \alpha + \beta T_l + \varepsilon_{il} \qquad (1)$$

where Y_{il} is the 2017 outcome for individual i from baseline locality l. T is an intent-to-treat (ITT) indicator that equals one for localities randomly assigned to T1 and zero for localities randomly assigned to T2, so that the estimate of β gives the impact of 18-month earlier exposure to Progresa. Given random assignment, our main specifications only controls for the 5 regions and 4 village-population size groups through fixed effects. Standard errors are adjusted for clustering at the locality level. For each cohort, we show estimates for the full sample, as well as separate estimates for women and men, and report whether treatment effects are significantly different across genders. For the early childhood cohort, we also show an alternative estimate that includes only children conceived before the start of the program (i.e., those born in in 1997 and 1998).

We first document differential impacts on educational attainment. For the early childhood outcome, educational attainment at age 18-20 is a primary outcome of interest, as it allows to test whether earlier investments in nutrition and health translated in long-term human capital gains (this cohort was too young to observe educational gains in the mid-term evaluation), as they start transitioning into the labor market. The other

[19] Authors´ calculations using the 2018 ENIGH survey, the National Income and Expenditure Household Survey. Data available on https://www.inegi.org.mx/programas/enigh/

primary outcome for this cohort is the income they expect to earn, which provides a way of gauging expected returns to the human capital investments once transition into adulthood is complete. For the school cohort, which at age 29-35 has fully transitioned into adulthood, educational gains are estimated to confirm that findings from the medium-term evaluations hold in the long term. For them, the primary outcomes are their income, geographic mobility, occupation, household formation, and fertility. Appendix E presents additional results to unpack some of the findings, as well as ITT effects on outcomes for which reliability was weaker and/or selection was a concern. For completeness, it also shows results for income, household formation and fertility of the early childhood cohort, but as these outcomes are incomplete and hence selected at age 18-20, they are harder to interpret and arguably observed too soon.

Results are robust to a large set of alternative estimations (available from the authors): i) we add baseline controls following Belloni et al (2014) post double selection lasso method;[20] ii) alternatively, we use no controls at all; iii) we only include observations where individuals were directly interviewed, excluding that from proxy informants;[21] iv) we limit the sample to the 13 strata without security problems; and v) we use village-level weights to overweigh villages in strata affected by insecurity so that the sample regains representativity of the initial 506 villages. These alternative specification and samples confirm the robustness of the results, though, we lose some precision when restricting to 13 of the 20 strata. Results are generally more precise when including only direct interviews, and when using the post double selection lasso controls.

5. Long-term differential impact estimates

5.1. The 20-year differential impact of Progresa for the early childhood cohort

Education

Table 4 and Figure 2 show that individuals with 18-month earlier exposure to Progresa in early childhood have higher educational achievement 20 years later (on average, 0.35 more grades attained). Differential educational gains are observed at all educational levels, including a 5 percentage point (p.p.) increase in the likelihood of completing lower secondary school (a 7 percent increase), a 4 p.p. increase in the likelihood of completing upper secondary school (a 18 percent increase), and a 2 p.p. increase in the likelihood of having completed already some tertiary education (an 65 percent increase, starting from a very low level of 3 percent in T2).

[20] Specifically, we follow Duflo (2018), starting from a large list of baseline controls (33) and including the squares and two-way interactions of all variables, and creating an indicator variable for observations with missing baseline data, then replacing the missing variable with 0.

[21] In the main tables, we present results on the largest sample available, including answers from proxy respondents when available. As not all outcomes were asked to the proxy respondents, that explains the differences in number of observations between the tables.

Educational gains are observed for both genders but results for higher levels of education are larger in magnitude and more significant for women, for whom the probability of finishing upper secondary school increased by 22 percent and the probability of university studies doubled (compared to 3 percent in T2). Consistent with higher education levels, being assigned to T1 decreased the likelihood of finishing school in the village of origin. The last column confirms results are robust to the exclusion of children born in 1999.

These results constitute new evidence on the impacts of exposure to Progresa during early childhood. From past evidence we know that earlier exposure to the cash, conditions, and information from the program led to improved nutrition and to investments in preventive healthcare. Our results suggest this translated into educational gains 20 years later. Moreover, as young adults in T1 were more likely to still be studying in 2017 (though not significantly so), these results may be an underestimate of final educational gains.

Education results for this cohort are consistent with investments in early life having improved cognitive and/or socio-emotional outcomes, possibly leading to timely school entry and/or better school progression. Further evidence pointing in that direction comes from the 2003 mid-term evaluation data. Children born in 1997 in T1 villages were 6 p.p. more likely to have attained literacy at age 5 (according to parental report) compared to those born in T2 (with the mean being 23 percent). See Table C1 and related discussion in Appendix C. Overall, having benefitted from the transfers early in life gave children in this cohort an advantage relative to children for whom exposure started 18 months later.

Income expectations

To understand long-term returns, it would be important to know whether educational gains translated into income and welfare increases. This cohort was 18-20 years old in 2017 and many were still studying so we observed them too soon in their adult life to answer this question.[22] Instead we collected information on these individuals´ expectations regarding their income at age 30 (the approximate age of the school cohort in 2017). As the income expectation variable is understandably noisy, we present a log and a trimmed version of the estimation (rank or winsorized versions show similar results). [23]

Table 5 shows that, consistent with their higher levels of education, the individuals in T1 have significantly higher income expectations (8 percent) than those in T2. Point estimates are substantially larger for girls (13

[22] Table E2 shows that 55 percent of them is active in the labor market, with a small but not significant positive differences between T1 and T2. Given that, the positive and significant point estimates for labor income shown in the same table can reflect both an extensive and intensive margin effect but is not necessarily a meaningful estimate of long-term income gains. Similarly, for these 18-20 year old, results on fertility and marriage are still censored so only presented for completeness in Table E3.

[23] Income expectations can be a hard question to answer, and 12.5 percent of the individuals (15 percent of girls and 9 percent of boys) indicated they could not answer this question. Table 5 shows that non-response is balanced between T1 and T2.

percent*)*, consistent with their higher differential education outcomes. Results for boys are not significant. Results are slightly stronger when excluding individuals born in 1999.

Expectations on education complement this analysis. Table 5 shows that individuals in T1 are more likely (at 10 percent significance level) to expect finishing tertiary education than those in T2. Interestingly, more than half of the individuals in both groups expects to finish tertiary education. Hence, even if only 22 percent were enrolled at the time of the survey, many expect to re-enroll, a finding similar to that by Duflo, Dupas and Kremer (2019) in Ghana. This confirms that this cohort´s transition into adulthood is incomplete.

Appendix E replicates all of the results discussed above with separate treatment effects for individuals born in 1997, 1998 and 1999, and tests the differences between birthyears. The purpose of this estimation is to explore whether in the early childhood cohort, the impact of differential exposure to Progresa differed within periods of the first 1000-day window. We find no significant differences between birthyears in either educational gains or income expectations but recognize this could be due to statistical power or to the fact that this cohort has yet to complete its education (Table E1).[24]

5.2. The 20-year differential impact of Progresa for the school cohort

Education

Tables 6 to 11 show the 20-year differential effects for the cohort transitioning from primary into secondary school during the 18 months when only T1 received transfers. Focusing first on educational achievement in Table 6, and consistent with results from the medium-term evaluation, 20-years later there is still a relatively large and significant differential (7 p.p., a 12 percent increase compared to T2) in the likelihood of having completed lower secondary school. There is no differential in having finished primary school, almost universal for this cohort, nor in having completed higher levels of education. This suggests that differential exposure helped a substantial share of children in T1 complete lower secondary school who, otherwise, may have dropped out after primary. It did not, however, help them achieve levels beyond lower-secondary. This is intuitive, as the group that otherwise would have dropped out of primary school was unlikely to progress beyond lower secondary. Hence, the differential timing of the start of the transfer shifted this cohort up one level of education, but not more (Figure 3). On average, this translated in about 0.3 more grades attained, a difference that is likely to be the permanent differential given this cohort´s age and their low likelihood of still

[24] Point estimates of the differential treatment effect for having completed lower secondary school (the highest level that all three cohorts could have finished given their age), are larger for those born in 1998 or 99 (0.061 and 0.069 compared to 0.038), but not significantly so.

being in school. Results are similar when only considering individuals directly interviewed, though slightly stronger (consistent with the measures obtained by proxy being noisier). [25]

The ITT estimates on finishing lower secondary school are larger for girls (10 p.p.) than for boys (4 p.p.). For other education outcomes, point estimates for girls are also substantially larger than for boys, but the difference is not significant in most cases. The same pattern is observed for most of the final outcome variables we consider next. The gender difference in finishing lower secondary school mirrors that found in earlier work, with the point estimates being very close to the short-term differences in enrollment in secondary school for this cohort, as shown in Schultz (2004).

Did increased schooling improve income in the long-term? There are several reasons this could be the case. There could be returns to education, because completing lower secondary school may make it easier to access certain jobs, or because the skills acquired may increase productivity and therefore wages and income.[26] In addition, going to lower secondary school also meant that this cohort was more likely to have completed its schooling outside of the village of origin, and this result is particularly strong for girls (14 p.p.). As such, schooling mechanically increased students´ mobility which could have helped them expand their horizons, and possibly their readiness to migrate. Moreover, going to school in a different town would have exposed them to new and possibly better-connected networks. The latter could be relevant given the importance of networks for (international) migration in Mexico.

Income

Table 7 estimates impacts on labor force participation and labor income, over the past 12 months and from the two main occupations. Assignment into T1 during the transition from primary to lower secondary school results in a relatively large and significant gain in annual labor income by 2017. In line with the educational differential, effects are larger for women. Notably, there is no differential effect on labor force participation, which is almost universal among men in this cohort, but as low as 38 percent among women. However, there is a significant difference in the probability of having some labor income in US dollars. The probability of having income originating in the US increases from 4 percent in T2 to 8 percent in T1. There are clear gender differences in

[25] The long-term differential estimates on educational attainment in Table 6 are a bit smaller than the medium-term results (Table C1) for 2003, when around one third of the children in this cohort were still enrolled in school. This is consistent with T2 possibly staying in school a little longer to complete (with a delay) lower secondary education. The point estimate of the differential impact of Progresa on enrollment in 2003 is indeed negative (Table C1). A comparison of T2 means between 2003 and 2017 suggests that another 10 p.p. of children in the school cohort completed lower secondary schooling after 2003, with children on average gaining about an additional year of education since 2003. Note, however, that the comparison between 2003 and 2017 comes with a caveat, as the 2003 survey has higher attrition rates.
[26] Results from the medium-term evaluations have raised questions on whether such learning gains were effectively achieved in Mexico (Behrman, Parker and Todd, 2009a). One particular type of learning that may be relevant for our results could be that completing lower secondary school also implied gaining some basic knowledge of English.

the magnitude of this effect although impacts are significant for both. Women are less likely to earn incomes in US dollars, but the magnitude of the impact is larger in relative terms. From a T2 mean of 2 percent we see an increase of 2.8 p.p. resulting from differential exposure to Progresa. For men, the magnitude of the impact is smaller in relative terms, but still considerable, a 74 percent increase over a T2 mean of 8 percent.

The US income can help explain the magnitude of the overall income gains, as average earnings are much higher in the US than in Mexico. Results are presented in nominal value, using PPP adjustments for US residents, and trimming outliers (winsorizing gives similar results). Given that there are no differences on the extensive margin, we also show income for those working (in logs) to reflect returns to education closer to wages for similarity with a large part of the literature. With the PPP adjustment, the differential impact on income is 16 percent, and 25 percent for women.

Only a small group of individuals received US incomes in 2017, raising questions on how much of the income results are driven by this group. Migration to the US is affected by T1 so we cannot merely compare migrants' incomes to those of non-migrants due to selection. Figure 4 depicts the income distribution of all individuals in the sample. It shows that the gains in income broadly occur in the top quartile or quintile (with the test of equality of the distributions significant at the 10 percent level). The quantile regressions in Table 8 confirm this finding. Given that 8 percent of those assigned to T1 had US income, the effects on income are *also* coming from domestic income. To interpret the skewed nature of these gains, it helps to note that in the school cohort the differential timing of exposure affected mostly a subgroup who would have dropped out after primary school in absence of the transfers.

Table 8 presents impacts on asset mobility and broadly confirms this pattern. The quantile regressions show a gradual shift upwards only at higher percentiles. The average effect on the asset index is positive but not significant.

Geographic mobility

Table 9 shows that both internal and international migration were frequent amongst individuals assigned to T1, who were 5 p.p. less likely to live in their village of origin in 2017. In fact, on average, they lived 55 km further away from it.[27] Increased US migration seems to be offset by a decrease in migration to large metropolitan areas in Mexico, in particular for men. In contrast, both men and women assigned to T1 were more likely to migrate

[27] Migrant destinations were asked down to the locality level, and distances were calculated using GPS coordinates of the localities. In cases where the specific locality names were missing, location was approximated using the GPS coordinates of the main town of the municipality where the individual lived in 2017. For US locations, often only state names are reported, in which case distances were calculated using common destination cities or the geographic center of the states.

to semi-urban or rural municipalities in Mexico and, on average, they lived in smaller municipalities (with less population).

We further note that the share of individuals that ever migrated to the US is much higher than those currently in the US – pointing to relatively large return migration. Return migration is indeed a well-known phenomenon in Mexico. The share of people having ever migrated is also 4 p.p. higher in T1 than in T2 (significant when only considering direct responses). Other variables indicating domestic migration (ever migrated out of municipality and whether left and then returned to the municipality) are also significantly higher, further confirming the overall larger mobility of T1.

These changes in mobility patterns are relatively large, and possibly consistent with people moving towards locations where there are higher returns to education or to their networks. Completing lower secondary education may not have allowed them to be competitive in urban labor markets, nor necessarily to shift into different occupations, but it may have made a difference in smaller towns. It could also have helped people realize where there could be higher returns to their skills, elsewhere in Mexico or in the US.

Table 10 (bottom line) confirms that few individuals were working in professional or technical jobs (4 percent), so the gains in income were not driven by a shift to skill-intensive jobs. This is consistent with the low share of individuals in this cohort that finished upper secondary school. Overall, we find few significant differences in occupations. Results hence suggest that the impacts of differential exposure to Progresa on labor income come from moving towards better-paying jobs in similar occupations. When we distinguish between wage employment or self-employment (something we can only do for individuals interviewed directly), there does appear to be a shift away from non-agricultural wage work for men, and into self-employment in agriculture or commerce. These are occupations that need some initial investment in assets, which possibly was facilitated by US income from return migrants (Table E4). Given the many outcomes tested for occupation, and the fact that this is only significant at the 10% and for men only, it needs to be interpreted with caution. The survey directly conducted with the individuals in the sample also allows to estimate ITT impacts on a few other indicators capturing characteristics of their jobs, including frequency of pay, regularity, and social security benefits. Results in Table E5 confirm that there is no evidence of people moving to better-quality jobs.

The broader pattern of the findings that emerges is that, even in a labor market with many frictions and average low returns, completing lower secondary education allowed people to find higher returns through geographic mobility, rather than through occupational mobility.

Household formation and fertility

Did differences in educational attainment translate into decisions on household formation or fertility? About 80 percent of the men and women in the school cohort were married and had children in 2017, and there are no significant differences between T1 and T2 on these extensive margin indicators (top two lines in Table 11). This facilitates the interpretation of the change in the age of marriage and age at the birth of the first child. Individuals assigned to T1 significantly postponed marriage and first childbearing with half a year, as compared to those in T2, and this is observed amongst women and men, with no significant differences between them (line 3 and 4 in Table 11). This is further illustrated by the difference in the cumulative distribution functions in Figure 5 and Figure 6, which show a shift to the right, with differences in childbearing starting around age 18 and increasing up to about ages 22-23.

As Figure 5 shows, this shift does not result from a reduction in teenage pregnancies, and hence is unlikely to be capturing an incarceration effect, often found when childbearing starts right after school dropout so that programs inducing girls to stay in school lead to fewer early pregnancies. Instead, individuals in T1 waited an additional half a year in their early twenties before committing to household and parenting responsibilities. It is plausibly indicative of a shift in strategy, with a first period when individuals migrated to earn income, to later settle back in (more rural settings in) Mexico.

The shift in age of first birth also implies that at age 30, women in T1 were more likely (compares to T2) to still have very young children under their care. Table 11 shows they are more likely to have children under 2, and the average of the youngest child is also smaller. This possibly provides an explanation as to why despite them achieving higher education, women's labor force participation at age 30 was not higher. At the same time, there were no significant differences between T1 and T2 in the total number of children (2 on average), and no differences in birth spacing. Hence, in terms of long-term fertility outcomes, there may be few differences. However, the delayed start of childbearing and household formation, possibility helped these individuals to settle in a location where the returns to their education were higher or to accumulate some wealth before starting a family. This is important to understand the overall welfare implications of exposure to Progresa.[28]

Treatment heterogeneity: presence of a lower secondary school in the village

We hypothesized that school-related mobility could have played a role in migration decision and extended networks. It can therefore be informative to analyze treatment heterogeneity based on whether children would

[28] Understanding whether the program affected welfare through matching in the marriage market is less straightforward. While we asked basic information on partners´ education, age, and occupation, it is often incomplete, partly because household proxy informants would not know this information. Even when reported, it suffers from more measurement error (Appendix D reports results on the reliability of proxy measures). With those caveats, we note that we find little significant differences in partner characteristics when they were reported (Table E6). In line with the overall education results, we do find that the T1 individuals are more likely to have higher education than their partners, as compared to T2, possibly improving their bargaining position in the household.

have needed to commute outside of their village of origin to attend lower secondary school. The cash transfers could have helped cover transportation costs, and more so for girls who received higher amounts (and for whom the risks of walking long distances may have been perceived to be higher).

We explore if impacts differ depending on whether or not there was a lower secondary school in the village of origin in 1997. About 20 percent of the villages had one. Table E7 presents these results for the main outcome variables. Villages with lower secondary schools present at baseline have higher educational attainment, higher income and higher mobility, but being exposed to treatment early did not significantly move any of these indicators (none of the P-values in the bottom line are significant). The point estimates of the treatment effects on income and mobility in villages without secondary school presence at baseline are higher than in the full sample (though not significantly so), which is consistent with school-related mobility and networks as a potential mechanism to explain long-term results. But as the treatment effect on completing secondary education is also larger for this group, these results do not allow to separate out the mobility or network mechanisms from other effects of education.

5.3. **Comparison between the early childhood and the school cohorts**

The focus of this paper is on two cohorts for whom the 18-months difference in exposure between T1 and T2 falls at possible critical stages in childhood. The rationale for the selection of each of those cohorts is different, and the ITT estimates may capture different parameters. This is the case, in part, because ITT estimates of the school cohort, by design, are more likely to be driven by the "marginal" children (i.e. those who, in absence of the transfer, would have dropped out after primary school) while the ITT estimates of the early childhood cohort capture exposure to the nutrition and health components of the program earlier in life, hypothesized to have led to general health and cognitive gains. The ITT estimates of the early childhood cohort may also capture dynamic complementarities between exposure during early childhood and during school age, if those were larger for those assigned to T1.

The results hence do not necessarily provide evidence on whether the impact of Progresa varies with the age at which children were exposed to it. Notwithstanding, we note that assignment into T1 in the early childhood cohort shifts the entire education distribution to the right (Figure 2), consistent with the hypothesized general cognitive gains, which contrasts with the shift in lower secondary school found for the school cohort (Figure 3).

Geographic mobility is harder to compare because the early childhood cohort has not yet completed its schooling. That said, the one clear contrast between the two cohorts is in international migration. Fifteen percent of the school cohort has international migration experience, with an additional 5 p.p. differential treatment effect, and three-fourths of it occurred prior to age 20.[29] In contrast, only 3 percent of the early childhood cohort migrated internationally, with no differential impacts on this outcome (Table 12). These differences reflect more general changes in migration flows from Mexico (and from the 7 states in the sample) to the US, which reached a peak in 2007, when the school cohort was around 20 years old and declined very sharply since then (Figure 7). A decline in the frequency of undocumented immigration from Mexico to the US over this period has been show by Massey, Durand and Pren (2015) and Parrado and Ocampo (2019). Table 12 further allows to compare geographic mobility of the early childhood cohort within Mexico with those of the school cohort (in Table 9). Given that 22 percent of the younger cohort is still studying, mobility likely in part directly reflects school locations, given the absence of institutions offering higher levels of education in the program localities and municipalities. Results confirm that T1 individuals are more likely to have migrated and more likely to have gotten their last schooling outside of the municipality, compared to T2. And possibly consistent with gains at higher levels of education, there is a differential impact on migration to urban areas (outside of the metropolitan areas), rather than the trend towards rural or semi-rural destinations observed for the school cohort.

For both cohorts, we find women's income (or income expectations) to be lower than men's, despite the fact that women on average achieved higher schooling. Both cohorts exhibit differential impacts on education and income (or income expectations) that are stronger for women. This is consistent with earlier exposure to the CCT helping women of both cohorts to compensate for other factors that limited their income-earning potential.

Comparing the income expectations result of the early childhood cohort with the income result of the school cohort shows the expectations about returns to education are broadly consistent with actual returns of the older cohort. Interestingly, the absolute value of the magnitude of the differential treatment effect for expected income is in line with that of actual income among the school cohort. A speculative explanation could be that the larger education levels across the board for the early childhood cohort help compensate for the reduced international migration opportunities, which resulted in income gains for the school cohort.[30]

6. Siblings

[29] Based on retrospective data on the age of the first international migration, asked the 2017 survey. Only one person of the school cohort reported having his last schooling in the US. Children in this cohort hence finished schooling and migrated to the US shortly afterwards. Among the early childhood cohort, 16 persons reported having gone to school in the US, which likely captures children who migrated with their parents at young ages.

[30] Mean income expectations amongst T2 in the early childhood cohort (Table 5) are substantially larger than actual incomes of 30 year-olds (Table 7) in the school cohort. Expectations and actual incomes for men are closer in magnitude. This is because all women (and men) reported positive expected incomes, while in the older cohort only 38 percent of women participate in an income-earning activity. This suggest that women reported expected income under the assumption they would be working at age 30, caveating the comparison made.

By design, our survey leaves unanswered the question of whether there were differential impacts of exposure to Progresa on income 20 years later for cohorts other than the two discussed in the paper. A possible reason for long-term differentials on other cohorts is that T1 households, by virtue of starting earlier, also would have received a larger accumulated total amount of cash over the 20-year period, something that may have benefitted all children.[31] On the other hand, impacts for the early childhood and school cohorts may have been driven by households pooling resources away from other siblings towards those who benefitted from the fortuitous timing of Progresa´s rollout.

To investigate these questions, we collected information on educational attainment and geographic mobility for all other members who were children at baseline from the households in our sample, henceforth referred to as the siblings sample.[32] Information was either obtained during the proxy interview with the Progresa beneficiary of the household (typically the mother of the children for whom information was collected), or during the interview with the individuals in our two cohorts (who reported proxy information about their siblings).[33] Results are not representative of all children in eligible households, as they are restricted to children with siblings in our two cohorts. Even so, the differential treatment effects can still provide suggestive evidence. Moreover, with the 2003 data we can further benchmark results for this particular sample relative to the full sample of children (the 2003 survey attempted to collect information on all children). Appendix C presents this analysis (Tables C2-C5), as well as further details on the 2017 sibling sample estimations

We explore impacts on siblings by grouping them based on their age and on their baseline schooling. Results in Table C6 show that the differential exposure to Progresa led to smaller but still significant long-term educational gains for 9-15-year-old siblings who had been at lower grade levels (first panel) or not enrolled at all (3rd panel), compared to those of individuals transitioning from primary to lower secondary school (the main sample). For these siblings, who were either lagging behind or had dropped out entirely, the differential exposure led to an 8 p.p. increase in the probability of finishing primary school. The higher probability of finishing primary school did not translate into more geographic mobility to any destination. As primary schools were located in the villages of origin, primary school attendance did not facilitate geographic mobility or expand networks. Skills learned in primary school are likely to be less relevant for mobility.

[31] Appendix A documents these differences in transfer amounts and shows that take-up was marginally (but significantly) higher in T1 than in T2.

[32] To be precise, this sample does not strictly only include siblings, but all children (some of which were cousins, for instance) who were younger than 18 years of age in 1997, and who lived in eligible households where there was at least one child in our two cohorts.

[33] To reduce burden on the respondents, field protocols were designed to avoid duplication of information (i.e., proxy information on an individual was intended to be only collected from one member of the original household). Even so, such duplicate information was collected for 18 percent of siblings (e.g., when members of the same households of origin were being interviewed by different survey teams at the same time). For cases with multiple information, we average across answers of the different respondents, thereby likely reducing measurement error in the proxy report.

Among the 9-15-year-old siblings who had finished 6th grade by the start of the program (2nd panel), and among 16-17-year-old siblings (4th panel), we find no significant differential effects of earlier exposure to Progresa on their education, or their mobility.[34] This suggests that higher mobility of the school and early childhood cohorts did not affect that of their siblings and also that intra-sibling trade-offs within households were limited (or possibly that both factors cancel each other out on average). As we do not find increased mobility among the siblings of the main cohorts of interest (indeed many of the coefficients are negative) these results also suggest that the increased mobility of the main cohorts is not primarily driven by the difference in transfer amounts between T1 and T2, further pointing to the a plausible relationship between human capital and geographic mobility.

Considering the long-term effects on siblings in the age group between our two main cohorts (i.e. 1 to 8 year olds at baseline) shows that by 2017, there was a significant difference in years of education attained between T1 and T2 (Table A7).[35] Point estimates are smaller than for the early childhood cohort, but gains are present beyond the lower secondary level, driven by those between ages 1-5 at the start of Progresa. Together with the results by birth year shown in Table E1, these findings show that children generally benefitted from earlier exposure to Progresa anytime between conception and age 5, which are the ages typically included in early childhood interventions. The results further suggest that children who started receiving Progresa earlier benefited more, though differences are not significant. Even so, the qualitative finding is arguably interesting especially since the difference in the cumulative amount of transfers received by households in T1 compared to T2 is smaller since younger children were more likely to live in households that did not get the education grants during the experimental variation period (e.g. the probability of having a sibling in the school cohort was 30 percent for the 1-2 year old siblings, and 47 percent for the 6-8 year old siblings).

7. Conclusion

More than 20 years ago, Mexico was among the very first countries in the world to test an innovative model of social protection, with a redistributive, short-term goal of reducing poverty and inequality, while at the same time investing in the human capital of the next generation. It also was among the first countries to embrace evidence-based policy making building a large-scale and rigorous evaluation in the initial program design, which was instrumental to establish credible evidence of the short- and medium-term impacts of the program. These two innovations help explain why the Mexican conditional cash transfer model was later adopted by many other countries worldwide.

---

[34] Admittedly, sample sizes are smaller for these last two groups, so we may lack statistical power to detect small differences.
[35] Behrman, Parker and Todd (2009b) also showed that those who were 1-2 years old in 1997 entered school at a younger age.

After more than 20 years of experience with conditional cash transfers, it is important to understand their long-term impacts. Did the short-term gains in human capital investments translate in sustainable improvements in education, labor market and life outcomes? This paper builds on the initial rigorous evaluation design of the first phase of Progresa in rural areas of seven central states of Mexico, to analyze whether experimental variation in exposure during critical ages in childhood led to long-term differential outcomes. Although such differential outcomes likely provide an underestimate of the total absolute effects, they can offer important proof-of-concept regarding the existence of long-term effects of the program. Finding such differences, as we do in this paper, offers strong evidence that human capital investments facilitated by cash transfer programs are translated in improved outcomes in the beneficiaries' lives 20 years later. The paper further highlights that understanding the long-term effects necessarily requires accounting for mobility of the beneficiaries, most of whom 20 years later have left their original households, often indeed moving far away from their villages.

Amongst individuals randomly assigned to Progresa in early childhood, having been exposed to Progresa 18 months earlier in life resulted in 0.4 years more schooling by the time they were 18-20 years old. Differential exposure to the program during early childhood led to an 8 percent increase in the likelihood of finishing secondary school, an 18 percent increase in that of finishing upper secondary school, and a 67 percent increase in obtaining university studies. Overall, earlier exposure led to a shift of the educational attainment distribution at all levels, with results being particularly strong for women. Additionally, and consistent with the educational outcomes, differential exposure to Progresa during early childhood increased expected labor income, especially for women. It also increased geographic mobility.

The cohort randomly exposed to the program when they were about to make the transition from primary to secondary school, rather than 18 months later, experienced an increase of 7 p.p. in the likelihood of finishing lower secondary school (9th grade) – with effects for women larger at 10 p.p. Earlier exposure translated into a 15 percent increase in annual labor income - 25 percent for women, and more geographic mobility. It increased the likelihood of international migration and of earning income abroad. Internal migration shifted towards small urban and semi-urban areas. Income gains are likely to come from beneficiaries' ability to move to places with higher paying jobs, including in self-employment, because we found no significant changes in the probability of being economically active or in the type of occupation. These gains were concentrated on a subset of individuals, suggesting others did not overcome frictions in the labor market. Lastly, earlier exposure to the program delayed parenthood and household formation with about half a year (from a control mean of 21 years), but it did not change the probability of having children.

When evaluating the returns to education for Progresa beneficiaries, it is important to place their educational gains in the national context. Educational levels increased a lot in Mexico over this period, and for the same cohorts. Average schooling was higher at the national level than among individuals from the evaluation villages (consistent with the targeting of the program to marginal localities and poor households). Hence, any returns to education occur in a context where many others -with higher educational levels- were competing for jobs in the labor market. This is not unlike the situation of many other programs that increase schooling of poor children across the world, given the large gains in educational attainment in many low and middle-income countries. The higher income and shifts in trajectories may result not only from signaling or increased skills coming from more schooling but could also be explained by access to different networks and from attending school outside of the village which opened the path of geographic mobility. This too is likely to be relevant in many other low- and middle-income country settings.

In sum, this paper shows that conditional cash transfers in Mexico contributed to important gains in, and returns to human capital, both through exposure to the health and nutrition components very early in life, and through exposure to the educational component during the transition from primary to lower secondary school. This finding is notable given that in 2019 the program was substantially transformed, eliminating the health and nutrition component of Prospera, and focusing grants on upper-secondary and tertiary education, while giving a more modest amount to families with children enrolled in preschool, primary or lower-secondary school.

More generally, the evidence in this paper is unique in showing experimental impacts at-scale 20 years after the start of nation-wide government program, and it does so for a social program that has been replicated across Latin America and many other parts of the world. These positive findings arguably re-emphasize the value of the initial innovations of the CCT approach, especially as the differential results are likely an underestimate of the total absolute effects. At the same time, the results reveal important heterogeneity in outcomes: while a subset of beneficiaries was able to use the opportunities provided by the program to substantially improve educational and labor market outcomes (including through international migration), others may have benefitted less. These results both confirm the fundamental strengths of the initial CCT approach, but also point to a need for complementary policies to allow more households to bear the full fruit of increased opportunities. Finally, the results on early childhood exposure not only highlight the returns to the nutrition and health components of the CCT, but more generally provide unique large-scale long-term evidence on the returns to investments in human capital in the early years of life.

References

Almond, D., J. Currie, and V. Duque. 2018. "Child Circumstances and Adult Outcomes: Act II." *Journal of Economic Literature* 56(4): 1360–1446.

Angelucci, M, 2015. "Migration and financial constraints: evidence from Mexico", *The Review of Economics and Statistics***,** 97(1).

Araujo, M. C., M. Bosch, and N. Schady. 2018. "Can Cash Transfers Help Households Escape an Intergenerational Poverty Trap?" in C. Barrett, M.R. Carter and JP Chavas (eds.). *The Economics of Poverty Traps*, University of Chicago Press.

Attanasio, O. and K. Kaufmann, 2014. "Education choices and returns to schooling: Mothers' and youths' subjective expectations and their role by gender", *Journal of Development Economics*, 109, 203-216

Baird, S., C. McIntosh, and B. Özler. 2019. "When the money runs out: Do cash transfers have sustained effects on human capital accumulation?", *Journal of Development Economics*,140:169-185.

Baird, S., JH Hicks, M. Kremer, and E. Miguel, 2016. "Worms at work: Long-run Impacts of a Child Health Investment", *The Quarterly Journal of Economics* 131 (4), 1637-1680

Banerjee, A., E. Duflo, and G. Sharma. 2021. "Long-term Effects of the Targeting the Ultra Poor Program" *American Economic Review: Insights*, forthcoming.

Barham, T. 2011. "A healthier start: The effect of conditional cash transfers on neonatal and infant mortality in rural Mexico," *Journal of Development Economics*, 94(1),74-85.

Barham, T., K. Macours, and J.A. Maluccio. 2013. "Boys' Cognitive Skill Formation and Physical Growth: Long-Term Experimental Evidence on Critical Ages for Early Childhood Interventions," *American Economic Review: Papers & Proceedings* 103(3): 467–471.

Barham, T., K. Macours, and J.A. Maluccio. 2018a. "Are Conditional Cash Transfers Fulfilling Their Promise? Schooling, Learning, and Earnings after Ten Years" *CEPR Discussion Paper* 11937.

Barham, T., K. Macours, and J.A. Maluccio. 2018b. "Experimental Evidence of Exposure to a Conditional Cash Transfer During Early Teenage Years: Young Women's Fertility and Labor Market Outcomes" *CEPR Discussion Paper* 13165.

Baseler, T. 2021. "Hidden Income and the Perceived Returns to Migration: Experimental Evidence from Kenya". Mimeo, University of Rochester. Available at SSRN: https://ssrn.com/abstract=3534715.

Bazzi, S., A. Gaduh, A. D. Rothenberg, and M. Wong. 2016. "Skill Transferability, Migration, and Development: Evidence from Population Resettlement in Indonesia." *American Economic Review*, 106 (9): 2658-98.

Beegle, K., J. De Weerdt and S. Dercon, 2011. "Migration and Economic Mobility in Tanzania: Evidence from a Tracking Survey," *The Review of Economics and Statistics*, 93(3): 1010-1033.

Behrman, J. R., S. W. Parker, and P. E. Todd. 2009b. "Schooling Impacts of Conditional Cash Transfers on Young Children: Evidence from Mexico." *Economic Development and Cultural Change* 57 (3): 439–77.

Behrman, J.R. and Hoddinott, J. 2005. "Programme Evaluation with Unobserved Heterogeneity and Selective Implementation: The Mexican *PROGRESA* Impact on Child Nutrition." *Oxford Bulletin of Economics and Statistics*, 67: 547-569.

Behrman, J.R., S. Parker, and P.E. Todd. 2009a. "Medium Term Impacts of the Oportunidades Conditional Cash Transfer Program on Rural Youth in Mexico", in Stephan Klasen and Felicitas Nowak-Lehmann, Eds., *Poverty, Inequality and Policy in Latin America, Cambridge* , MA: MIT Press.

Behrman, J.R., S. Parker, and P.E. Todd. 2011. "Do Conditional Cash Transfers for Schooling Generate Lasting Benefits? A Five-Year Followup of PROGRESA/Oportunidades." *The Journal of Human Resources*, 46(1):93–122.

Belloni, A., V. Chernozhukov, and C. Hansen. 2014. "Inference on treatment effects after selection among high-dimensional controls." *Review of Economic Studies* 81: 608–650.

Bettinger, E., M. Kremer, M. Kugler, C. Medina, C. Posso, and J.E. Saavedra. 2016. "Can Educational Voucher Programs Pay for Themselves?" mimeo, Stanford University.

Blattman, C., N. Fiala and S. Martinez. 2020. "The long term impacts of grants on poverty: 9-year evidence from the Youth Opportunities Program in Uganda". *American Economic Review: Insights. 2020,* 2(3): 287–304

Bobba, M. and Gignoux, J. 2019. "Neighborhood Effects in Integrated Social Policies". *World Bank Economic Review*, 33:1, 116-139.

Bryan, G. and M. Morten. 2019. "The Aggregate Productivity Effects of Internal Migration: Evidence from Indonesia". *Journal of Political Economy*.127:5, 2229-2268.

Bryan, G., S. Chowdhury and M. Mobarak. 2014. "Under-investment in a Profitable Technology: The Case of Seasonal Migration in Bangladesh". *Econometrica*, 82(5): 1671–1748.

Cahyadi, N., R. Hanna, B.A. Olken et al. 2020. "Cumulative Impacts of Conditional Cash Transfer Programs: Experimental Evidence from Indonesia." *American Economic Journal: Economic Policy*, 12(4):88-110.

Caponi, V. 2011. "Intergenerational Transmission of Abilities and Self-selection of Mexican Immigrants". *International Economic Review*, 52: 523-547.

Chiquiar, D. and G.H. Hanson. 2005. "International Migration, Self-Selection, and the Distribution of Wages: Evidence from Mexico and the United States." *Journal of Political Economy*, vol. 113(2): 239–281.

Davis, B., Stecklov, G., and Winters, P. 2002. "Domestic and International Migration from Rural Mexico: Disaggregating the Effects of Network Structure and Composition." *Population Studies*, 56 (3): 291–309.

Donato KM, Hamilton ER, and Bernard-Sasges A. 2019. "Gender and Health in Mexico: Differences between Returned Migrants and Nonmigrants". *The ANNALS of the American Academy of Political and Social Science*. 684(1):165-187.

Duflo, E., 2018. "Machinistas meet Randomistas: Useful ML tools for empirical researchers", *NBER Summer Institute lecture*. http://conference.nber.org/conf_papers/f114791.slides.pdf

Duflo, E., P. Dupas and M. Kremer. 2015 "Education, HIV, and Early Fertility: Experimental Evidence from Kenya." *American Economic Review* 105(9): 2757–2797.

Duflo, E., P. Dupas and M. Kremer. 2021. "The Impact of Free Secondary Education: Experimental Evidence from Ghana. *NBER Working Paper 28937*.

Fernald Lia C., P. Gertler, and L. Neufeld. 2009. "10-year effect of Oportunidades, Mexico's conditional cash transfer programme, on child growth, cognition, language, and behaviour: a longitudinal follow-up study". *Lancet*. Dec 12;374(9706):1997-2005.

Gertler, P. 2004. "Do Conditional Cash Transfers Improve Child Health? Evidence from PROGRESA's Control Randomized Experiment." *American Economic Review*, 94 (2): 336-341.

Gertler, P. J. Heckman, R. Pinto, A. Zanolini, C. Vermeersch, S. Walker, S.M. Chang, and S. Grantham-McGregor. 2014. "Labor Market Returns to an Early Childhood Stimulation Intervention in Jamaica." *Science* 344(6187) 997–1001.

Gertler, P. J. Heckman, R. Pinto, S.M. Chang, S. Grantham-McGregor, C. Vermeersch, S. Walker, and A. Wright. 2021. "Effect of the Jamaica Early Childhood Stimulation Intervention on Labor Market Outcomes at Age 31." NBER *working paper 29292*.

Gertler, P., S. Martinez, and M. Rubio-Codina. 2012. "Investing Cash Transfers to Raise Long Term Living Standards." *American Economic Journal: Applied Economics*. 4(1): 164-92.

Gollin, D., D. Lagakos, and M. E. Waugh. 2014. "The Agricultural Productivity Gap", *The Quarterly Journal of Economics*, 129(2):39–993

Hamory, J., E. Miguel, M. Walker, M. Kremer, and S. Baird. 2020. "Twenty Year Economic Impacts of Deworming". *Proceedings of the National Academy of Sciences*, 118 (14) e2023185118

Hamory, J., M. Kleemans, N. Y Li, and E. Miguel. 2020. "Reevaluating Agricultural Productivity Gaps with Longitudinal Microdata", *Journal of the European Economic Association*; 19(3): 1522-1555.

Heckman, J. J. 2006. "Skill formation and the economics of investing in disadvantaged children." *Science*, 312(5782):1900–1902.

Hoddinott, J., J.A. Maluccio, J.R. Behrman, R. Flores, and R. Martorell. 2008. "Effect of a Nutrition Intervention During Early Childhood on Economic Productivity in Guatemalan Adults." *Lancet* 371(9610, 2): 411–416.

Ibarraran, P., N. Medellin, F. Regalia, and M. Stampini (editors). 2017. *How conditional cash transfers work: good practices after 20 years of implementation*, Washington DC: Inter-American Development Bank.

Kugler, A. and Rojas, I. 2018. "Do CCTs Improve Employment and Earnings in the Very Long-Term? Evidence from Mexico". *NBER Working Paper 24248*.

Lagakos, D. 2020. "Urban-Rural Gaps in the Developing World: Does Internal Migration Offer Opportunities?" *Journal of Economic Perspectives*, 34 (3): 174-92.

Lagakos, D., S. Marshall , A.M. Mobarak , C. Vernot, and M.E. Waugh. 2020. "Migration costs and observational returns to migration in the developing world*, Journal of Monetary Economics*, 113L: 138-154.

Levy, S. 2018. *Under-Rewarded Efforts: The Elusive Quest for Prosperity in Mexico*, Inter-American Development Bank

Li, S. 2018. "Investment and interruption: effects of the US experience on the earnings of return migrants in Mexico", *Applied Economics*, 50:4, 426-440.

Lindstrom, D.P. 1996. "Economic Opportunity in Mexico and Return Migration from the United States." *Demography*, 33(3):357–374.

Massey DS, Durand J, and Pren KA. 2015. "Border Enforcement and Return Migration by Documented and Undocumented Mexicans". *J Ethn Migr Stud*. 41(7):1015-1040.

McKenzie, D. and H. Rapoport. 2007. "Network effects and the dynamics of migration and inequality: Theory and evidence from Mexico". *Journal of Development Economics*, 84(1): 1-24.

McKenzie, D. and H. Rapoport. 2010. "Self-selection patterns in Mexico-US migration: the role of migration networks." *The Review of Economics and Statistics*, 92(4):811–821.

McKenzie, D., J. Gibson and S. Stillman._2010. "How Important is Selection? Experimental Vs Non-experimental Measures of the Income Gains from Migration ", *Journal of the European Economic Association*, 8(4): 913-45.

Molina-Millan, T., K. Macours, J. Maluccio, and L. Tejerina, 2020. "Experimental Long-Term Effects of Early Childhood and School-Age Exposure to a Conditional Cash Transfer Program" *Journal of Development Economics*, 143(102385)

Molina-Millan, T., T. Barham, K. Macours, J. Maluccio, and M. Stampini. 2019. "Long-Term Impacts of Conditional Cash Transfers: Review of the Evidence." *World Bank Research Observer*, 3(1): 2019, 34(1): 119–59.

Munshi, K. 2011. “Strength in Numbers: Networks as a Solution to Occupational Traps”, *The Review of Economic Studies*, 78(3):1069–1101.

Munshi, K. 2020. “Social Networks and Migration”. *Annual Review of Economics*. V12: 503-524.

Parker, S. and P. Todd, 2017. “Conditional Cash Transfers: The Case of Progresa/Oportunidades”, *Journal of Economic Literature* 55 (3), 866-915.

Parker, S. and Vogl, T. 2018. “Do Conditional Cash Transfers Improve Economic Outcomes in the Next Generation? Evidence from Mexico”. *NBER Working Paper 24303*.

Parrado EA, and Ocampo AN. 2019. “Continuities and Changes in the Processes of Mexican Migration and Return”. *The ANNALS of the American Academy of Political and Social Science*. 684(1):212-226.

Progresa. 1997. “Nota Tecnica: Diseno de muestras basal y control”, mimeo.

Rivera JA, Sotres-Alvarez D, Habicht JP, Shamah T, and Villalpando S. 2004. “Impact of the Mexican program for education, health, and nutrition (Progresa) on rates of growth and anemia in infants and young children: a randomized effectiveness study”. *JAMA*, 291: 2563–70. 9.

Schultz, T.P. 2004. “School subsidies for the poor: evaluating the Mexican Progresa poverty program”. *Journal of Development Economics*. 74(1): 199-250.

Skoufias, E., 2005. “Progresa and Its Impacts on the Welfare of Rural Households in Mexico”, *IFPRI Research Report* 139. Washington DC.

Stecklov G, Winters P, Todd, and J, Regalia F. 2007. “Unintended effects of poverty programmes on childbearing in less developed countries: experimental evidence from Latin America.” *Popul Stud* (Camb). 61(2):125-40.

Stecklov, G P Winters, M Stampini, and B Davis, 2005. “Do conditional cash transfers influence migration? A study using experimental data from the Mexican PROGRESA program”, *Demography* 42 (4), 769-790

Ullmann, S.H, N. Goldman, and D. S. Massey. 2011. “Healthier before they migrate, less healthy when they return? The health of returned migrants in Mexico”, *Social Science & Medicine,* 73(3):421-428.

Waldman K, Wang JS, and Oh H. 2019. “Psychiatric problems among returned migrants in Mexico: updated findings from the Mexican Migration Project”. *Soc Psychiatry Psychiatr Epidemiol.* 54(10):1285-1294.

Wilson, F.A. J.P. Stimpson, and J.A. Pagán. 2014. "Disparities in Health Outcomes of Return Migrants in Mexico", *International Journal of Population Research*, vol. 2014,

Winters, P., A. de Janvry and E. Sadoulet (2001) “Family and Community Networks in Mexico-US Migration”, *Journal of Human Resources* 36(1): 159-84.

World Bank, 2018. *Learning to Realize Education’s Promise.* World Development Report 2018. Washington DC.

Young, A. 2013. “Inequality, the Urban-Rural Gap, and Migration”, *The Quarterly Journal of Economics*, 128(4): 1727–1785.

Figures and Tables

Figure 1: Villages of origin and locations in 2017 of individuals in the tracking sample

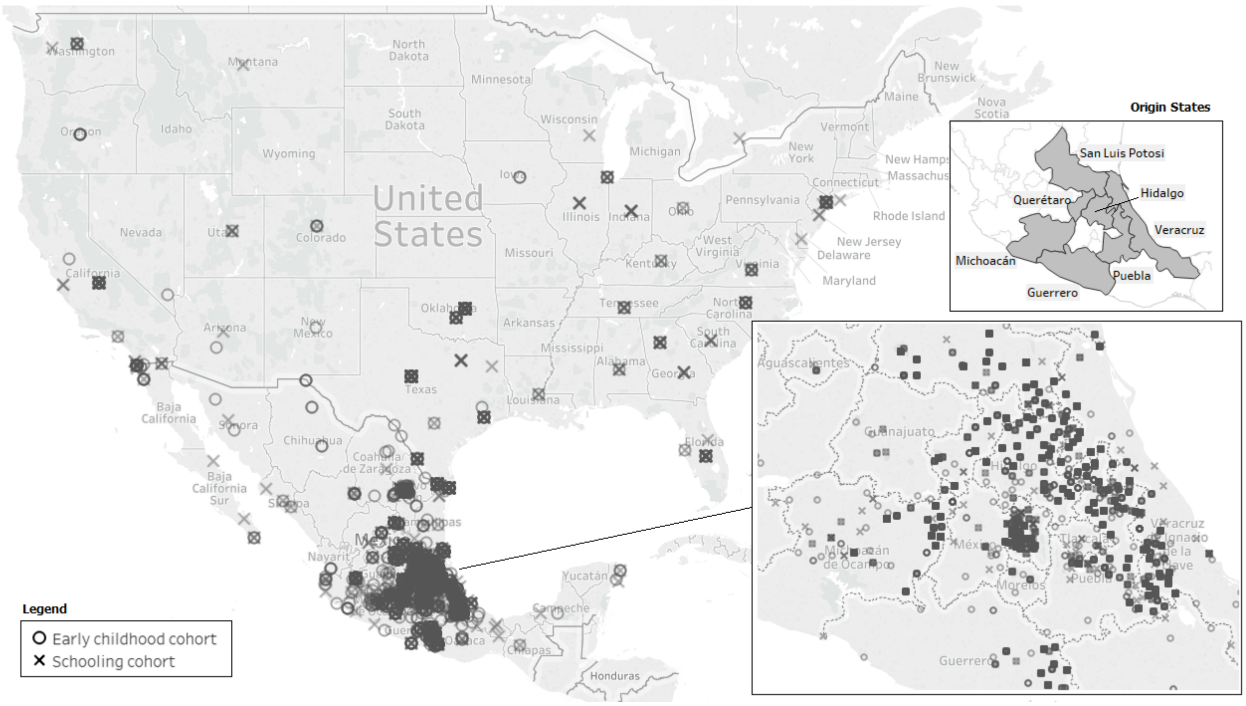


Figure 2: Cumulative distribution of grades attained in 2017 - early childhood cohort

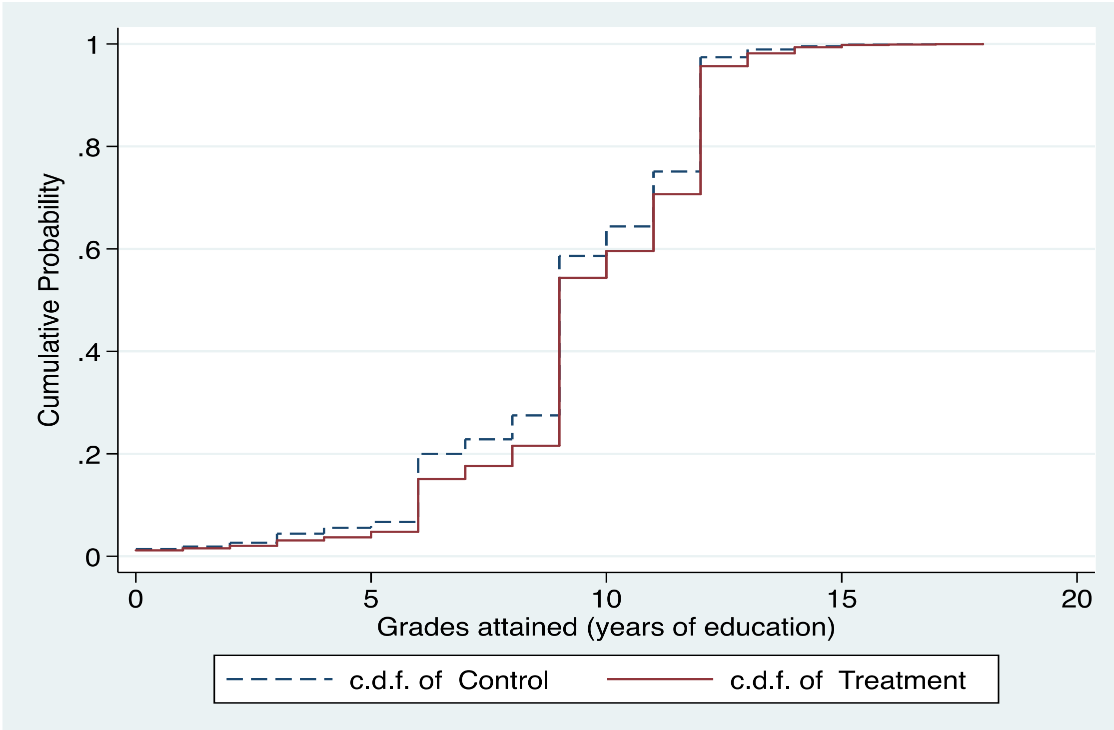


Figure 3: Cumulative distribution of grades attained in 2017 - school cohort

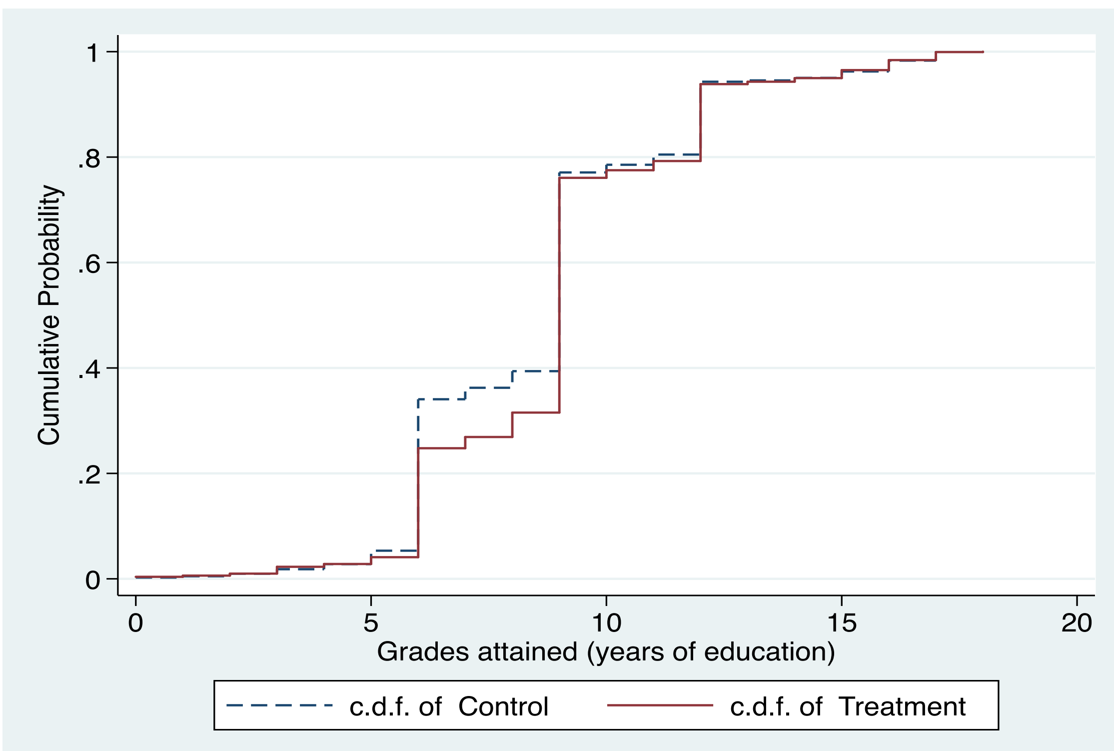
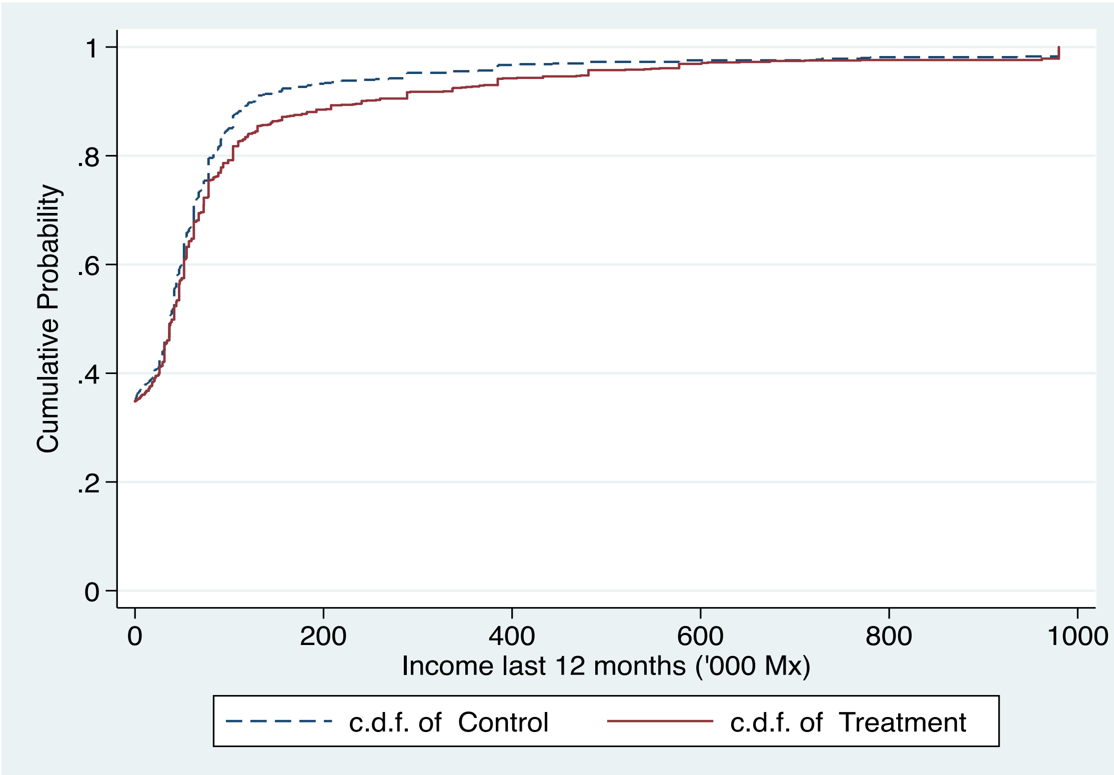


Figure 4: Cumulative distribution income – school cohort



Note: Cumulative Income Distribution. P-value Ksmirnov test of equality of distributions is 0.087.

Figure 5: Cumulative distribution of age at which school cohort women had their first child

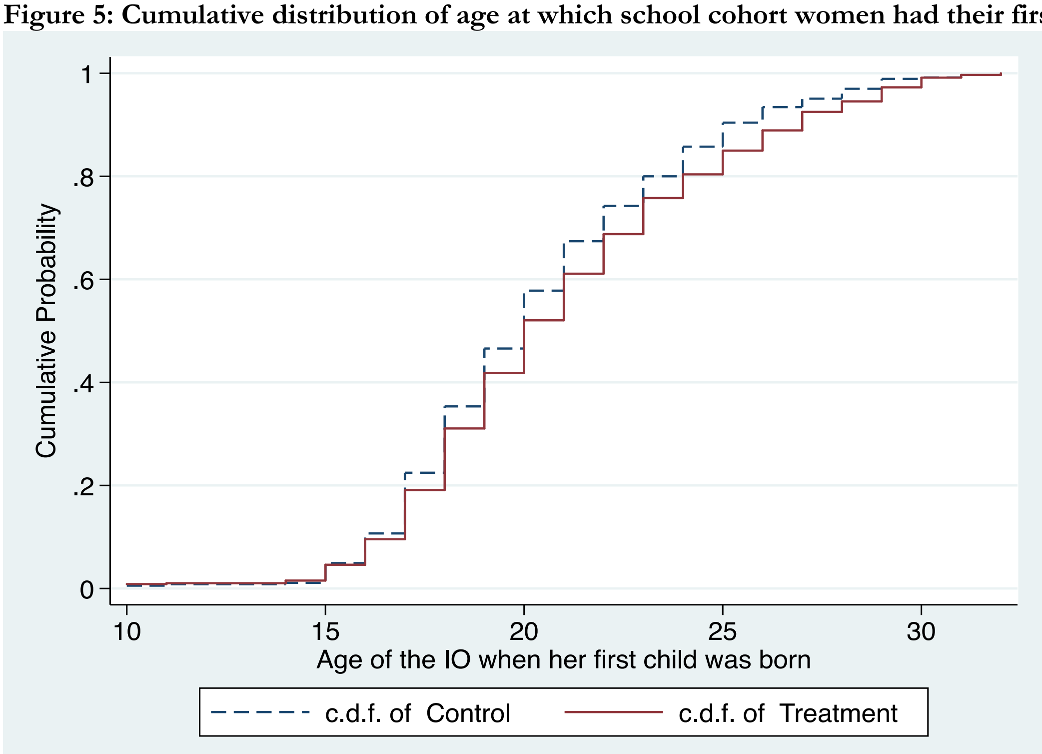


Figure 6: Cumulative distribution of age at which school cohort men had their first child

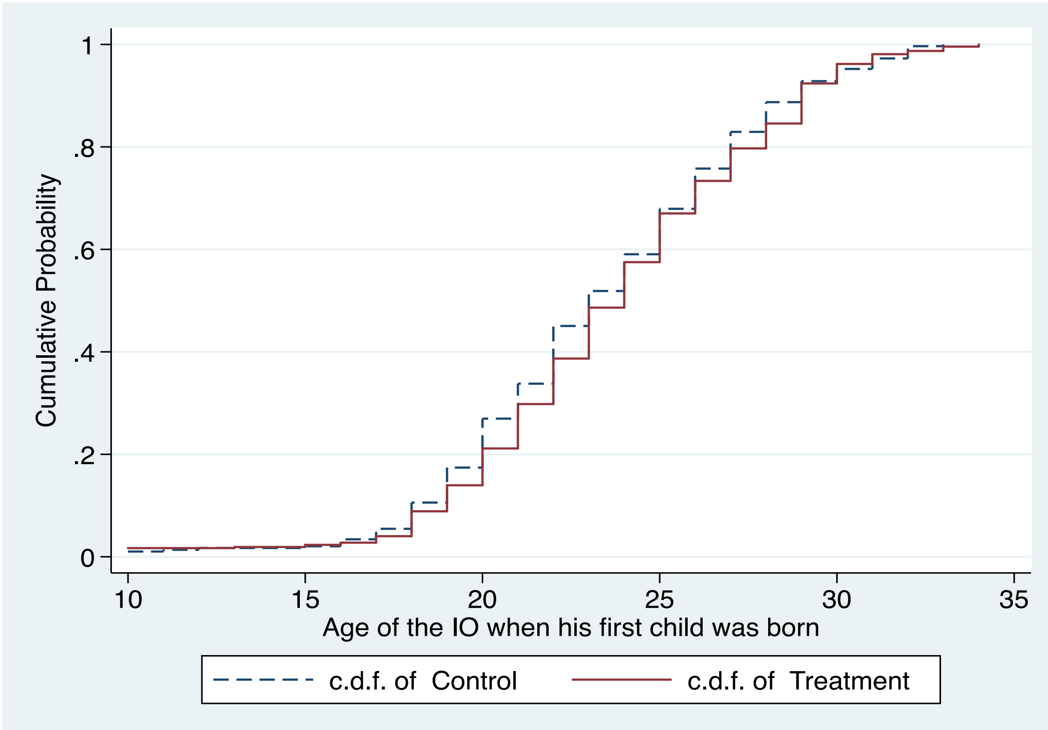


Figure 7: Migration flows (number of migrants leaving Mexico for the US, per year)

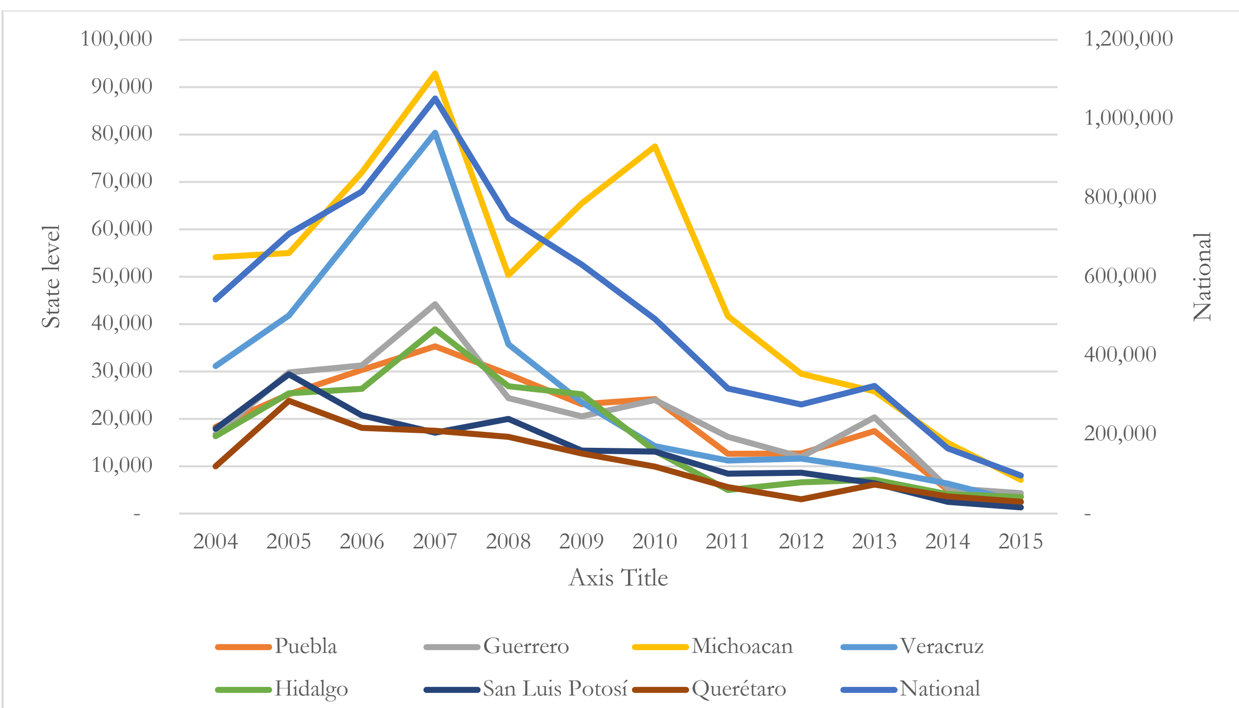


Table 1: Tracking rates and balance by type of survey (or outcome) and balance test

	Face-to-face, by phone or proxy	Face-to-face or by phone	Face-to-face	Death
School cohort (N=2289)				
T1	-0.004	0.018	-0.037	-0.006
	(0.011)	(0.02)	(0.026)	(0.006)
mean T2	0.943	0.813	0.426	0.023
Early childhood cohort (N=4461)				
T1	0.007	0.004	-0.016	-0.006
	(0.01)	(0.013)	(0.02)	(0.005)
mean T2	0.923	0.857	0.567	0.028

Note: Tracking results after intensive tracking of full sample

Table 2: Place of residence of individuals in 2017

	School cohort			Early childhood cohort		
	N	% all	% found	N	% all	% found
USA	2,231	0.12	0.11	4,317	0.03	0.03
In Mexico, by type of location						
Metropolitan areas	2,231	0.18	0.18	4,317	0.14	0.13
Other urban municipality	2,221	0.28	0.28	4,306	0.28	0.28
Semi-urban municipality	2,221	0.30	0.31	4,306	0.37	0.39
Rural municipality	2,221	0.13	0.13	4,306	0.17	0.18
With respect to place of origin						
State of origin	2,245	0.60	0.61	4,350	0.75	0.77
Municipality of origin	2,245	0.47	0.48	4,350	0.63	0.66
Locality of origin	2,245	0.34	0.36	4,350	0.52	0.55

Note: Mexican destinations cover 30 different Mexican states (out of 32). US destinations include 28 US States + Canada (1 observation)

Table 3: Inter-generational educational mobility

	Household head		School cohort	Early childhood cohort
Progresa sample				
Year	1997	1997	2017	2017
Age	44	~ 30	~ 30	18-20
Grades attained	3	4.1	8.8	9.5
Primary completed (%)	24	41	95	94
Lower secondary completed (%)	4	8	65	78
Upper secondary completed (%)	<1	<1	20	35
Some tertiary education (%)			6	6
Currently studying (%)			2	23
Mexican national averages				
Year			2018	2018
Age			~ 30	18-20
Grades attained (nationally)			10.4	10.3
Grades attained (those living in rural areas)			8.5	9.7

Note: Estimates for top panel based on 1997 ENCASEH and 2017 tracking survey. Estimates for national averages in lower-panel based on 2018 ENIGH survey, the National Income and Expenditure Household Survey

Table 4: 20-year differential treatment effects for the early childhood cohort: educational attainment

	All		Women		Men		P-value	Born in 1997 or 1998	
	Obs	coef.	Obs	coef.	Obs	coef.	Women=Men	Obs	coef.
	(1)		(2)		(3)		(4)	(5)	
Completed primary school									
T1	4103	0.017	2069	0.015	2034	0.02	0.804	3009	0.018
s.e.		(0.01)*		(0.013)		(0.013)			(0.01)*
mean T2		0.933		0.943		0.923			0.931
Completed lower secondary school									
T1	4103	0.053	2069	0.051	2034	0.056	0.861	3009	0.048
s.e.		(0.024)**		(0.027)*		(0.027)**			(0.024)**
mean T2		0.725		0.751		0.697			0.731
Completed upper secondary school									
T1	4103	0.044	2069	0.055	2034	0.034	0.416	3009	0.026
s.e.		(0.02)**		(0.026)**		(0.024)			(0.024)
mean T2		0.249		0.262		0.235			0.3
Completed some tertiary education									
T1	4103	0.017	2069	0.028	2034	0.007	0.09*	3009	0.021
s.e.		(0.006)***		(0.009)***		(0.008)			(0.008)***
mean T2		0.026		0.028		0.023			0.032
Grades attained (years of education)									
T1	4103	0.354	2069	0.391	2034	0.324	0.688	3009	0.322
s.e.		(0.153)**		(0.178)**		(0.173)*			(0.161)**
mean T2		9.133		9.27		8.987			9.213
Last schooling in locality of origin									
T1	3760	-0.071	1934	-0.074	1826	-0.069	0.898	2734	-0.08
s.e.		(0.031)**		(0.036)**		(0.033)**			(0.031)**
mean T2		0.392		0.38		0.405			0.394
Still studying									
T1	4137	0.021	2083	0.029	2054	0.012	0.562	3036	0.042
s.e.		(0.019)		(0.023)		(0.024)			(0.017)**
mean T2		0.215		0.222		0.207			0.15

Notes: ITT estimates with region and village population size fixed effects. Col. 1 shows results for the full early childhood cohort, Col. 2 for female and Col. 3 for males. Col. 4 shows the p-values for tests that the effects are identical between males and females. Col (5) shows results restricting to those born in 97 and 98. Each column shows the number of observations and the estimated differential treatment effect in the first row, standard errors are in the second cell row parentheses, clustered at locality level, with ***, **, * indicating significance at 1, 5 and 10%; and mean in T2 localities in the third row.

Table 5: 20-year differential treatment effects for the early childhood cohort: income and education expectations

	All		**Women**		**Men**		**P-value**	**Born in 1997 or 1998**	
	Obs	coef.	Obs	coef.	Obs	coef.	Women=Men	Obs	coef.
	(1)		(2)		(3)		(4)	(5)	
Log (expected annual income at age 30)									
T1	3323	0.077	1655	0.129	1668	0.018	0.072*	2400	0.108
s.e.		(0.043)*		(0.058)**		(0.047)			(0.049)**
mean T2		11.654		11.571		11.741			11.622
Expected annual income at age 30 - trimmed at 2%									
T1	3260	11279.44	1618	14525.15	1642	6828.919	0.355	2359	13072.08
s.e.		(5474.95)**		(6866.928)**		(7410.115)			(6369.734)**
mean T2		144253.9		135129		153848.3			141273.4
Could/did not answer on income expected at age 30									
T1	3828	0.009	1967	0.008	1861	0.015	0.737	2781	0.01
s.e.		(0.017)		(0.023)		(0.018)			(0.018)
mean T2		0.125		0.152		0.093			0.128
Expects to finish tertiary education									
T1	3828	0.038	1967	0.039	1861	0.038	0.979	2781	0.033
s.e.		(0.022)*		(0.028)		(0.028)			(0.024)
mean T2		0.54		0.564		0.512			0.544

Notes: ITT estimates with region and village population size fixed effects. Col. 1 shows results for the full early childhood cohort, Col. 2 for female and Col. 3 for males. Col. 4 shows the p-values for tests that the effects are identical between males and females. Col (5) shows results restricting to those born in 97 and 98. Each column shows the number of observations and the estimated differential treatment effect in the first row, standard errors are in the second cell row parentheses, clustered at locality level, with ***, **, * indicating significance at 1, 5 and 10%; and mean in T2 localities in the third row.

Table 6: 20-year differential treatment effects for the school cohort: educational attainment

	All		Women		Men		P-value
	Obs	coef.	Obs	coef.	Obs	coef.	Women=Men
	(1)		(2)		(3)		(4)
Completed primary school							
T1	2141	0.013	1092	0	1049	0.026	0.216
s.e.		(0.011)		(0.016)		(0.016)	
mean T2		0.947		0.962		0.93	
Completed lower secondary school							
T1	2141	0.071	1092	0.103	1049	0.039	0.125
s.e.		(0.029)**		(0.035)***		(0.038)	
mean T2		0.606		0.615		0.597	
Completed upper secondary school							
T1	2141	0.012	1092	0.002	1049	0.023	0.536
s.e.		(0.023)		(0.029)		(0.029)	
mean T2		0.195		0.213		0.177	
Completed some tertiary education							
T1	2141	0.003	1092	0.007	1049	0	0.665
s.e.		(0.011)		(0.015)		(0.014)	
mean T2		0.057		0.064		0.05	
Grades attained (years of education)							
T1	2141	0.277	1092	0.337	1049	0.216	0.569
s.e.		(0.169)		(0.203)*		(0.219)	
mean T2		8.641		8.74		8.537	
Last schooling in locality of origin							
T1	1856	-0.101	992	-0.143	864	-0.054	0.051*
s.e.		(0.034)***		(0.041)***		(0.04)	
mean T2		0.5		0.525		0.47	
Still studying							
T1	2153	0.006	1097	0.006	1056	0.006	0.897
s.e.		(0.006)		(0.009)		(0.007)	
mean T2		0.015		0.019		0.01	

Notes: ITT estimates with region and village population size fixed effects. Col. 1 shows results for the schooling cohort, Col. 2 for female and Col. 3 for males. Col. 4 shows the p-values for tests that the effects are identical between males and females. Each column shows the number of observations and the estimated differential treatment effect in the first row, standard errors are in the second cell row parentheses, clustered at locality level, with ***, **, * indicating significance at 1, 5 and 10%; and mean in T2 localities in the third row.

Table 7: 20-year differential treatment effects for the school cohort: income

	All		Women		Men		P-value
	Obs	coef.	Obs	coef.	Obs	coef.	Women=Men
		(1)		(2)		(3)	(4)
Labor income last 12 months							
T1	1801	8761.301	978	7592.966	823	5586.758	0.826
s.e.		(5083.328)*		(4132.521)*		(8831.537)	
mean T2		57044.79		24241.18		98102.9	
Has US income							
T1	1834	0.041	987	0.024	847	0.056	0.236
s.e.		(0.012)***		(0.011)**		(0.023)**	
mean T2		0.044		0.016		0.08	
Has labor income							
T1	1834	0.000	987	-0.018	847	-0.013	0.854
s.e.		(0.024)		(0.032)		(0.012)	
mean T2		0.649		0.382		0.978	
Log(labor income last 12 months), conditional on working							
T1	1187	0.157	372	0.246	815	0.112	0.334
s.e.		(0.059)***		(0.106)**		(0.069)	
mean T2		11.018		10.757		11.143	

Notes: ITT estimates with region and village population size fixed effects. Income includes labor income in Mexican pesos of the 2 main occupations in the last 12 months, trimmed at 1%, and with dollar income of US residents converted to Mexican pesos using PPP-adjusted exchange rate. Col. 1 shows results for the schooling cohort, Col. 2 for female and Col. 3 for males. Col. 4 shows the p-values for tests that the effects are identical between males and females. Each column shows the number of observations and the estimated differential treatment effect in the first row, standard errors are in the second cell row parentheses, clustered at locality level, with ***, **, * indicating significance at 1, 5 and 10%; and mean in T2 localities in the third row.

Table 8: Quantile regressions of 20-year differential treatment effect for the school cohort: income and assets

Percentile	10	20	30	40	50	60	70	75	80	90
Labor income last 12 months										
T1	n.a.	n.a.	n.a.	0	3000	2447	2990	10400**	15600**	52000**
				(3496)	(3582)	(3309)	(4013)	(5236)	(7523)	(21955)
Asset index										
T1	0.000	0.000	0.000	0.000	0.000	0.047	0.053	0.091	0.180*	0.223
	(0.050)	(0.058)	(0.056)	(0.064)	(0.061)	(0.068)	(0.073)	(0.080)	(0.11)	(0.16)
Asset mobility										
T1	0.000	0.000	0.035	0.061	0.003	0.047	0.129*	0.187*	0.224**	0.267**
	(0.050)	(0.048)	(0.050)	(0.056)	(0.062)	(0.063)	(0.071)	(0.096)	(0.10)	(0.12)

Note: N=1854 (1825 for income). Income in Mexican pesos includes labor income of the 2 main occupations in the last 12 months, with dollar income of US residents converted to Mexican pesos using PPP-adjusted exchange rate. Asset index is first principal component of ownership of 10 assets in 2017. In the asset mobility estimates, the same outcome variable is used but the regression controls for baseline assets of the origin households. Standard errors, clustered by locality, in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 9: 20-year differential treatment effects for the school cohort: geographic mobility

	All		Women		Men		P-value
	Obs	coef.	Obs	coef.	Obs	coef.	Women=Men
	(1)		(2)		(3)		(4)
Living in locality of origin							
T1	2153	-0.049	1097	-0.07	1056	-0.03	0.269
s.e.		(0.028)*		(0.033)**		(0.034)	
mean T2		0.386		0.363		0.409	
Living in municipality of origin							
T1	2153	-0.014	1097	-0.035	1056	0.007	0.312
s.e.		(0.03)		(0.037)		(0.036)	
mean T2		0.487		0.509		0.464	
Ever migrated outside of the municipality of origin							
T1	1882	0.07	1002	0.088	880	0.044	0.367
s.e.		(0.029)**		(0.039)**		(0.033)	
mean T2		0.617		0.565		0.68	
Migrated and returned to municipality of origin							
T1	1882	0.037	1002	0.039	880	0.027	0.82
s.e.		(0.02)*		(0.023)*		(0.03)	
mean T2		0.189		0.136		0.255	
Ever migrated to the USA							
T1	2149	0.037	1094	0.032	1055	0.038	0.852
s.e.		(0.024)		(0.023)		(0.036)	
mean T2		0.201		0.097		0.31	
In USA in 2017							
T1	2149	0.032	1096	0.032	1053	0.03	0.923
s.e.		(0.018)*		(0.017)*		(0.026)	
mean T2		0.09		0.042		0.139	
Distance (km) from locality of origin to 2017 location							
T1	2149	55.008	1096	61.471	1053	45.421	0.724
s.e.		(38.732)		(36.197)*		(58.776)	
mean T2		299.976		185.583		420.629	
Log(Population size of the largest town in municipality if in Mexico in 2017)							
T1	1865	-0.33	1006	-0.198	859	-0.476	0.202
s.e.		(0.161)**		(0.198)		(0.197)**	
mean T2		10.374		10.325		10.431	
By type of destination in Mexico							
In metropolitan areas of Mexico in 2017							
T1	2149	-0.055	1096	-0.03	1053	-0.082	0.137
s.e.		(0.026)**		(0.031)		(0.033)**	
mean T2		0.207		0.191		0.224	
In other urban municipality of Mexico in 2017							
T1	2139	0.02	1089	0.02	1050	0.02	0.945
s.e.		(0.023)		(0.034)		(0.026)	
mean T2		0.189		0.224		0.152	
In semi-urban municipality of Mexico in 2017							
T1	2139	0.028	1089	0.024	1050	0.034	0.649
s.e.		(0.017)*		(0.025)		(0.018)*	
mean T2		0.101		0.138		0.062	
In rural municipality of Mexico in 2017							
T1	2139	0.025	1089	0.027	1050	0.023	0.849
s.e.		(0.01)**		(0.015)*		(0.01)**	
mean T2		0.024		0.036		0.012	

Notes: ITT estimates with region and village population size fixed effects. Col. 1 shows results for the schooling cohort, Col. 2 for female and Col. 3 for males. Col. 4 shows the p-values for tests that the effects are identical between males and females. Each column shows the number of observations and the estimated differential treatment effect in the first row, standard errors are in the second cell row parentheses, clustered at locality level, with ***, **, * indicating significance at 1, 5 and 10%; and mean in T2 localities in the third row.

Table 10: 20-year differential treatment effects for the school cohort: occupation

	All		Women		Men		P-value
	Obs	coef.	Obs	coef.	Obs	coef.	Women=Men
		(1)	(2)			(3)	(4)
Main occupation in unpaid domestic work							
T1	2152	-0.003	1096	0.003	1056	0.002	0.841
s.e.		(0.021)		(0.032)		(0.004)	
mean T2		0.297		0.578		0.002	
Main occupation in agriculture							
T1	2152	0.016	1096	-0.003	1056	0.03	0.422
s.e.		(0.019)		(0.01)		(0.034)	
mean T2		0.158		0.028		0.295	
Main occupation in commerce							
T1	2152	0.016	1096	0.013	1056	0.021	0.755
s.e.		(0.013)		(0.017)		(0.02)	
mean T2		0.086		0.068		0.104	
Main occupation in services							
T1	2152	-0.006	1096	0.007	1056	-0.019	0.394
s.e.		(0.011)		(0.007)		(0.021)	
mean T2		0.07		0.012		0.132	
Main occupation in crafts							
T1	2152	-0.013	1096	-0.01	1056	-0.024	0.539
s.e.		(0.017)		(0.012)		(0.03)	
mean T2		0.132		0.035		0.233	
Main occupation in manual work/elementary tasks							
T1	2152	-0.02	1096	-0.017	1056	-0.022	0.81
s.e.		(0.014)		(0.021)		(0.018)	
mean T2		0.105		0.134		0.074	
Main occupation in non-agrircultural skilled job							
T1	2152	0.008	1096	0	1056	0.016	0.624
s.e.		(0.013)		(0.016)		(0.021)	
mean T2		0.087		0.073		0.102	
Main occupation in professional or technical job							
T1	2152	0.001	1096	0.003	1056	-0.001	0.705
s.e.		(0.009)		(0.012)		(0.012)	
mean T2		0.042		0.047		0.037	

Note: ITT estimates with region and village population size fixed effects. Col. 1 shows results for the schooling cohort, Col. 2 for female and Col. 3 for males. Col. 4 shows the p-values for tests that the effects are identical between males and females. Each column shows the number of observations and the estimated differential treatment effect in the first row, standard errors are in the second cell row parentheses, clustered at locality level, with ***, **, * indicating significance at 1, 5 and 10%; and mean in T2 localities in the third row. Occupations classified following INEGI 9 occupational categories, and aggregating the first 3 categories (directors, professionals, techical staff and administrative staff) in the "professional or technical job" given low frequency.

Table 11: 20-year differential treatment effects for the school cohort: marriage and fertility outcomes

	All		**Women**		**Men**		**P-value**
	Obs	coef.	Obs	coef.	Obs	coef.	Women=Men
	(1)		(2)		(3)		(4)
Ever married or in civil union							
T1	2153	-0.004	1097	0.003	1056	-0.012	0.723
s.e.		(0.019)		(0.026)		(0.025)	
mean T2		0.781		0.781		0.782	
Has any children							
T1	2134	0.016	1091	0.029	1043	0.005	0.577
s.e.		(0.016)		(0.02)		(0.027)	
mean T2		0.813		0.866		0.756	
Age at first marriage							
T1	1723	0.489	906	0.271	817	0.696	0.349
s.e.		(0.223)**		(0.286)		(0.296)**	
mean T2		21.537		20.28		22.927	
Age when had first child							
T1	1717	0.499	951	0.531	766	0.397	0.628
s.e.		(0.227)**		(0.287)*		(0.325)	
mean T2		21.702		20.367		23.365	
Age youngest child							
T1	1735	-0.332	953	-0.312	782	-0.293	0.75
s.e.		(0.171)*		(0.225)		(0.255)	
mean T2		4.891		5.381		4.287	
Has a 0-2 year old child							
T1	2114	0.04	1082	0.054	1032	0.022	0.346
s.e.		(0.019)**		(0.025)**		(0.03)	
mean T2		0.238		0.206		0.272	
Is household head or spouse of household head							
T1	2128	0.034	1092	0.052	1036	0.016	0.318
s.e.		(0.021)		(0.029)*		(0.028)	
mean T2		0.745		0.755		0.735	

Notes: ITT estimates with region and village population size fixed effects. Col. 1 shows results for the schooling cohort, Col. 2 for female and Col. 3 for males. Col. 4 shows the p-values for tests that the effects are identical between males and females. Each column shows the number of observations and the estimated differential treatment effect in the first row, standard errors are in the second cell row parentheses, clustered at locality level, with ***, **, * indicating significance at 1, 5 and 10%; and mean in T2 localities in the third row.

Table 12: 20-year differential treatment effects for the early childhood cohort: geographic mobility

		All		**Women**		**Men**		**P-value**	**Born in 1997 or 1998**	
		Obs	coef.	Obs	coef.	Obs	coef.	Women=Men	Obs	coef.
		(1)		(2)		(3)		(4)	(5)	
Living in locality of origin										
	ITT	4137	-0.042	2083	-0.036	2054	-0.052	0.628	3036	-0.051
	s.e.		(0.023)*		(0.028)		(0.027)*			(0.027)*
	Mean control		0.572		0.534		0.614			0.549
Migrated or willing to migrate outside of municipality										
	ITT	3978	0.036	2008	0.041	1970	0.027	0.48	2993	0.029
	s.e.		(0.014)***		(0.019)**		(0.014)*			(0.014)**
	Mean control		0.857		0.815		0.901			0.87
Ever migrated outside of the municipality										
	ITT	3830	0.065	1967	0.052	1863	0.079	0.38	2781	0.052
	s.e.		(0.024)***		(0.029)*		(0.03)***			(0.028)*
	Mean control		0.419		0.429		0.408			0.463
Migrated and returned to municipality										
	ITT	3830	0.021	1967	0.018	1863	0.023	0.656	2781	0.011
	s.e.		(0.016)		(0.019)		(0.02)			(0.017)
	Mean control		0.162		0.154		0.17			0.174
Last schooling in municipality of origin										
	ITT	3740	-0.035	1921	-0.052	1819	-0.02	0.223	2716	-0.041
	s.e.		(0.02)*		(0.025)**		(0.025)			(0.021)*
	Mean control		0.818		0.813		0.824			0.812
Ever migrated to USA										
	ITT	4133	0.004	2081	0.005	2052	0.002	0.74	3033	0.003
	s.e.		(0.008)		(0.007)		(0.013)			(0.008)
	Mean control		0.034		0.013		0.056			0.037
In USA in 2017										
	ITT	4126	0.001	2078	0.002	2048	-0.002	0.711	3027	0
	s.e.		(0.007)		(0.006)		(0.011)			(0.008)
	Mean control		0.026		0.011		0.042			0.029
Distance (km) from locality of origin to 2017 location										
	ITT	4127	6.522	2078	0.581	2049	10.151	0.788	3028	-5.537
	s.e.		(20.145)		(22.965)		(26.377)			(23.178)
	Mean control		132.727		106.673		160.603			155.138
Log(Population size largest town in municipality if in Mexico 2017)										
	ITT	3955	-0.178	2024	-0.222	1931	-0.133	0.546	2893	-0.225
	s.e.		(0.141)		(0.161)		(0.158)			(0.151)
	Mean control		9.823		9.858		9.786			9.995
By type of destination in Mexico										
In metropolitan areas in Mexico in 2017										
	ITT	4126	-0.008	2078	-0.015	2048	-0.003	0.587	3027	-0.023
	s.e.		(0.019)		(0.021)		(0.021)			(0.022)
	Mean control		0.134		0.13		0.139			0.157
In other urban municipality in Mexico in 2017										
	ITT	4122	0.042	2074	0.031	2048	0.055	0.347	3024	0.049
	s.e.		(0.017)**		(0.021)		(0.02)***			(0.02)**
	Mean control		0.14		0.169		0.109			0.148
In semi-urban municipality in Mexico in 2017										
	ITT	4122	0.005	2074	0.018	2048	-0.005	0.205	3024	0.015
	s.e.		(0.013)		(0.018)		(0.014)			(0.014)
	Mean control		0.089		0.104		0.073			0.087
In rural municipality in Mexico in 2017										
	ITT	4122	0.004	2074	0.001	2048	0.007	0.538	3024	0.011
	s.e.		(0.011)		(0.014)		(0.01)			(0.01)
	Mean control		0.036		0.05		0.022			0.029

Notes: ITT estimates with region and village population size fixed effects. Col. 1 shows results for the early childhood cohort, Col. 2 for female and Col. 3 for males. Col. 4 shows the p-values for tests that the effects are identical between males and females. Each column shows the number of observations and the estimated differential treatment effect in the first row, standard errors are in the second cell row parentheses, clustered at locality level, with ***, **, * indicating significance at 1, 5 and 10%; and mean in T2 localities in the third row.

Online Appendices for

Education, Income and Mobility:

Experimental Impacts of Childhood Exposure to Progresa after 20 Years

By M. CARIDAD ARAUJO ADN KAREN MACOURS

October 2021

List of Appendices

Appendix A: Past and current program participation

Tables A1 and A2 show different indicators of past and current program participation and compliance. The tables combine information obtained from the 2017 survey (hence referring to the self-reported participation as remembered and reported by the individuals), with information from the program's registry (both 2017 and historic administrative records), administrative information regarding transfer payments during early years (1998-2000, obtained through personal communication with Emanuel Skoufias) and during the period 1998-2003, compiled for previously published work (Gertler et al., 2012).

As payment of transfers and program status are the direct result of households' take-up and compliance decisions, they reflect the endogenous responses of households to program exposure, with the start of the timing of the latter being exogenously determined through the randomized assignments. As such, the data on past and current program participation helps to disentangle the mechanisms underlying the long-term effects.

School cohort

Information from administrative records

For the school cohort, administrative records document payments (combing health, nutrition, and education transfers) to households for the 1998-2003 period to 95 percent of T1 households and 93 percent of T2 households (Table A1). Hence, program-take up was high in both treatment groups with a small difference significant at the 10 percent level. As expected, given the earlier start of the program, households in T1 had accumulated more transfers by 2003.[1] It is likely that the absolute difference observed by 2003 provides a good estimate of the absolute difference over the 20-year period, as a large share of both types of households continued in the program after 2003 and this difference is, by construction, a relatively smaller share of total transfer received by 2017).

By 2017, 59 percent of the original households were still program beneficiaries, even if they were no longer receiving education transfers for the school cohort themselves.[2] Among the individuals of the school cohort, 23 percent were reported to still live in that original household receiving benefits in 2017 based on program

[1] For the cumulative transfer amounts between 1998 and 2003, the replication data of Gertler et al (2012) is used. By 2003, T1 had received the program for 5 years, and T2 for 3.5 years. As this dataset does not contain information on 11 percent of the target households, it was complemented with transfer data for the May 1998 to March 2000 period, obtained through personal communication with Emanuel Skoufias, in order to determine which households received any transfers. Using cumulative transfers up to 2000 from Skoufias, similarly, shows a large and significant difference in cumulative transfers for this earlier period between T1 and T2. We employed this approach because the data from the program registry (*padrón*) was incomplete for the first program years.

[2] After the initial determination of eligibility, there was a recertification process (in theory every 3 years) in which households were reinterviewed to determine if they continued to be eligible for the program. At that point, their beneficiary status was either renewed, or they transitioned into a scheme of partial benefits (called *Esquema Diferenciado de Apoyos* or EDA*)*, which included only secondary and upper secondary education grants but excluded primary-school grants and cash transfers associated with the health/nutrition components. Households that had left the program entirely, on average, stopped receiving transfers 12 years earlier (i.e., *circa* 2005).

administration data. Therefore, most were no longer considered household members for the purpose of attendance to health check-ups or information sessions. Almost none of the school cohort individuals (3 percent) was a direct beneficiary of the program in 2017 (defined as being either the direct recipient or the household head of a recipient household). There are no significant differences for these 2017 outcomes from the administrative records between treatment groups.

Information from self-report

We consider next what the target individuals themselves report about program participation. As program communication and payments were made to the main program beneficiary (typically the mother or grandmother of the individuals tracked in 2017), the target individuals themselves may not necessarily have been fully aware of these benefits, in particular if they did not receive school transfers. The self-report on program participation is also likely to suffer from recall error, given the long time passed since the beginning of the program. The reported program participation (take-up) is indeed lower than that obtained from the administrative data. Nevertheless, and acknowledging possible recall errors, the differences between treatment groups are still insightful.

For instance, for the school cohort, Table A1 shows that individuals from T1 are 14 p.p. more likely to report having benefitted from the program, and 19 p.p. more likely to report having received the school transfers specifically (from a mean of 63 percent in T2). The differences between treatment groups are 10 p.p. larger for girls than for boys. The number of years individuals report having benefitted from the school transfer is a full year longer in T1 than in T2 (1.1 year for girls and 0.8 of a year for boys). Individuals further report they were, on average, 11 years old when their families started receiving school transfers for them, and 15 years old when their school transfers stopped (with 22 percent reporting still receiving school transfers at age 18). Conditional on receiving transfers, ages of both first and last transfers are a slightly lower in T1. Overall, the results are consistent with an important share of individuals in T2 dropping out after completing primary school, making them ineligible for transfers once the program reached this group, and with T1 (in particular T1 girls) having benefitted more from the schooling transfers than those in T2, confirming results in earlier work.

With regard to program status in 2017, 36 percent report living in a household receiving program benefits in 2017, though in half of the cases this is a different household than that of 1997. That could possibly explain the higher share of reported program participation compared to that in the administrative data.

Early childhood cohort

Information from administrative records

For the early childhood cohort, administrative records show relatively high take-up that is marginally stronger in T1 (91 and 89 percent, Table A2). Consistent with 18-months delay in the start of the program for T2, cumulative transfers received by 2003 were significantly larger from those of T1 (30 percent). The total amounts are lower than those of the school cohort, as expected, given that transfer amounts are linked to the household demographics, with children that go to school receiving education transfers that increase by grade level.

Among the original households of the early childhood cohort, 61 percent were still program beneficiaries in 2017. While we cannot fully trace those born in 1998 and 1999 in the administrative data, of those born in 1997, 46 percent is listed as active member of a beneficiary household in the 2017 administrative database. Virtually none (less than 1 percent) are themselves direct program beneficiaries. There are no significant differences between treatment groups in these various 2017 program status variables.

Results from administrative data are broadly similar for both genders and across birthyears. However, there are no significant differences between treatment groups in take-up for the 1998 cohort.

Information from self-report

The self-reported household-level take up is similar to that from administrative data, and also shows a small (2 p.p.) but significantly higher take-up for T1. Interestingly, T1 is also slightly more likely to report having benefited from the schooling transfers, and – conditional on receiving them - benefited from them for an additional half a year (compared to 6 years for T2). This is consistent with T1 individuals of this cohort staying longer in school and obtaining higher education levels. For girls, this appears to be in part because they started receiving the schooling transfer at an earlier age, consistent with early enrollment and/or faster progression through the early grades of primary school.

By 2017, 62 percent of individuals reports living in a household benefitting from the program, and for most of them this is their original household. The 2017 status is not different between T1 and T2.

Appendix B: Sample design

We only consider households that were deemed eligible for the program according to their baseline proxy means score. The identification of the young cohort was further based on cross-validation of the information in the different rounds of ENCEL collected between 1998 and 2000. More specifically, we included children born in baseline households that were classified as poor, identified as infants before the 2000 survey and whose ages were: i) at most 0 years in 1997; ii) at most 1 year in 1998; or iii) at most 2 years in 1999. We also include those identified as age 1-3 in the 2000 surveys if they only appeared in one of the 2000 surveys and not before. We excluded i) individuals from households that were added to the ENCEL surveys after 1998, and for which no baseline and no poverty score was available; ii) those who had entry (entries) with age 0 in one of the ENCEL surveys, but who had inconsistent birthdates or age information in multiple other rounds of the survey; and iii) those included in the roster but marked as "non-existent", which are children who typically do not have birthdates nor relationships with the household head – and for whom only the first name was known.

For the early childhood cohort, this led to 4,677 observations. The tracking exercise revealed that 164 individuals identified as possible respondents based on the ENCEL surveys were duplicates of other observations. In addition, 52 individuals with very incomplete names and baseline information were determined to be erroneously included in the sample, after careful verification with household members. Those individuals are hence not included in the analysis, leading to a final sample of 4,461 individuals.

Surveys were completed for 4,137 individuals of the early childhood cohort. Note that among those there are 20 observations with birthdates seemingly before 97 that were included in the sample because they were identified in one of the ENCEL surveys as being younger, even if the 1997 ENCASEH contradicted that information (i.e., had an earlier birth dates). Only for 3 individuals the 2017 follow-up survey confirmed a birthdate in 1997. In addition, there were also some individuals in the early childhood cohort who, in 2017, reported birthyears that were before 1997 or after 1999, hence contradicting the information from the ENCASEH or early ENCEL waves. While these observations may be adding some noise to the estimates, we kept them in the sample, as we do not have similar information for those that attrited.

Identification of the children in the school cohort was done based on data from the 1997 ENCASEH, from which we selected all individuals in eligible households that were in primary school and had completed 5th grade by the time of the November 1997 ENCASEH (2,112 children). Because baseline data is missing for some households in the ENCEL panel, we also add children enrolled in 6th grade at the moment of the March 1998 ENCEL (192 children). Results are robust to exclusion of this 2nd group.

Of the total of 2,304 individuals identified for tracking in the school cohort, 15 individuals were identified during tracking as duplicated with other possible respondents, leading to the final sample of 2,289 individuals. Surveys were completed for 2,153 of them.

Appendix C: Medium-term (2003) results on the tracking sample, other cohorts and siblings

To allow for a cleaner comparison between Progresa´s medium-term impacts and our long-term impacts, we replicated the medium-term analysis with the 2003 data but restricting the sample to that of the tracking survey. We used the publicly available panel constructed by the program´s evaluation team, that links individuals of all the ENCELs. The top panel shows results for the 506 villages in the panel, and the bottom panel shows results for the tracking sample only (excluding the 47 villages we had to exclude from the 2017 survey due to insecurity).[3] For both cohorts, results are very similar for the full sample and the tracking sample.

In 2003, children of the early childhood cohort were 4 to 6 years old and transiting into primary school. Education outcomes in 2003 were only collected for individuals older than 5 years of age, hence the 2003 data only allows looking at outcomes for children of this cohort who were born in 1997. Table C1 shows that virtually all children born in 1997 were enrolled in school by age 5 and there is no difference in enrollment between T1 and T2. Interestingly, however, there is a significant difference in parental-reported literacy in 2003, with T1 children born in 1997 being about 6 p.p. more likely to already know how to read and write, compared to 23 percent in T2.

For the school cohort, the differential timing of exposure had led to a 11 p.p. higher probability of finishing lower secondary school by 2003 but did not have a significant impact on higher levels of education.[4] The last column of Table C1 shows that by 2003 there was no difference in enrollment, though this variable suffers from more attrition, as it was not asked for children that were no longer members of the baseline households. Among those included, around 30 percent of children were still enrolled in school, so the 2003 data does not allow to estimate final schooling gains.

Medium-term results for siblings

This section explores if there were experimental impacts 20 years later for cohorts other than the early childhood and the school cohort studied in this paper. We explored these impacts using the siblings sample. We first used the 2003 data to replicate the medium-term impacts of other cohorts, and to analyze how medium-term impacts of the siblings of the cohorts of interest compared to those of the entire group of children with similar age and education levels (i.e. unconditional on having siblings in the target ages).[5] Table C1 shows the medium-term results for the main cohorts for comparison. The first panels of Table C2 and C3 show results for the same outcomes as Table C1, but only for individuals who, at baseline, were 9 to 15 year old, enrolled in school, but had not completed 5 years of education (i.e. those lagging behind the school cohort). The second

[3] Sample sizes in panel 2 are lower than for the same cohort in 2017 due to larger attrition rates in the 2003 ENCEL.
[4] Attrition for education variables in 2003 was 13 percent and balanced between T1 and T2.
[5] Attrition rates in 2003 are between 11 and 25 percent and balanced between treatment groups.

panels of the same tables show results for individuals enrolled in 1997 who had completed 6 years of education. For both of these groups, the experimental variation in timing of the cash transfers did not coincide with transition to lower secondary school. We also show results for individuals who in 1997 were 9-15 year-old and were not enrolled in school (third panel), as the CCT could have helped them re-enroll. Finally, in the fourth panel we report results for an older cohort who was 16-17-year-old in 1997 and for whom the CCT arguably came too late. Table C2 focuses on all eligible individuals in the ENCEL sample, while Table C3 refers only to siblings of individuals in our two target cohorts. Table C2 confirms that differential educational gains were not limited to children that were in transition to lower secondary school, but they were also present on both those behind and those ahead of them, even if they were smaller (top 2 panels). The impact on the probability to complete lower secondary school was about half for the other enrolled children, when compared with estimated impacts for the target sample. Differential timing of exposure also helped some of the 9-15-year olds that were not attending school in 1997 to re-enroll, increasing their probability to complete primary school with almost 8 p.p. and their probability of completing lower secondary school with 5 p.p.. For the 16-17-year-olds who were too old to benefit from the educational components of the program, there are no differential effects, as expected.

Table C3 shows that these results broadly hold for the subsample of children that were siblings of the individuals in our target cohorts, with possibly somewhat stronger results for siblings that were lagging behind in terms of grade completion for the group of siblings 9-15 with 4 or less grades completed in 1997, for whom we see an increase of 8 and 6 p.p. in primary and lower secondary school completion, respectively. We return to this below.

Finally, we use the 2003 data also to replicate the medium-term results for the cohorts age 1-8 in 1997. These are children older than the early childhood cohort analyzed in this paper, but younger than the school cohort. As before, we show results for the full sample of all eligible households (Table C4), and then restrict the analysis to children that were siblings of the cohorts of interest, for whom we also have educational outcomes in 2017 (Table C5). The top panel of Table C4 shows that by 2003, differential educational gains of these cohorts (then 7-14 years old) were limited. There are no differentials on literacy, including for the cohort just older than those born in 97 - the 1-2-year-olds in 1997, and 7–8-year-olds in 2003. This contrasts with the impact on literacy among those born in 1997 shown in Table C1. For those children that had been transitioning into primary school when exposed to differential timing of transfers (i.e., those 12 to 14 in 2003 and with enrollment rates by then at 84 percent, bottom panel), there was a small educational gain of 0.19 years of education.

In line with results for the older siblings who had at most grade 4 completed, Table C5 shows that the differential timing of exposure did lead to slightly larger differences in educational outcomes by 2003 among siblings of our target cohorts, compared to the full sample of similar age groups. Hence individuals living in households with siblings exposed during critical ages seem to have benefitted a bit more (compared to those without such siblings) from the differential timing, possibly due to positive intra-household spillover effects. The result could also just indicate that households with more children in small ranges of age groups benefitted more from early exposure, possibly because it helped them address binding resource constraints and improve outcomes of all children. Overall, the 2003 results suggest that differential effects for the two main cohorts studied in this paper did not come at the cost of educational gains of other siblings.

Long-term impacts for siblings

Tables C6 and C7 show the results obtained using information on siblings' educational attainment in 2017. As for the target sample, sample sizes for each cohort are slightly larger than in 2003, given higher tracking rates. Attrition is higher than for the main cohorts (between 11 and 13 percent for the different subgroups of siblings) but balanced between treatment and control. We show the same outcomes as for the main sample.

Considering first the siblings of similar or older ages than the main school cohort (Table C6), the 2017 results are very similar to the 2003 results, confirming that the differential exposure led to long-term educational gains for the 9-15-year-old siblings of our target cohorts who had been at lower grade levels (first panel) or not enrolled at all (third panel). For both groups of siblings who were lagging behind or had dropped out entirely, the differential exposure led to an 8 p.p. increase in the probability of finishing primary school. For those who had advanced beyond primary school (second panel) or who were older (fourth panel), no differential long-term gains were observed. The 16-17-year olds can be considered a placebo group for the differential treatment effects. Separate estimates for men and women show similar results.

With the caveat that the quality of information may be lower for the siblings sample (as information was often obtained from a sibling rather than from a parent), the higher probability of finishing primary school did not translate into higher geographic mobility, neither to destinations in Mexico nor to the US. This result is in line with our main findings. Primary schools were located in the villages where individuals lived so attending primary school did not have the same implications as secondary school in terms of increasing social networks or geographic mobility. Skills learned in primary school may also be less relevant for mobility. The lack of mobility results for those who finished sixth grade by the start of the program and for the oldest cohort, further suggests that higher mobility of the individuals in the main sample neither increased nor decreased that of their siblings, in line with the lack of differential educational gains, and also suggesting that intra-sibling trade-offs within households were limited.

The long-term effects on siblings in the age group between the early childhood and the school cohorts point in the same direction. Table C7 shows that by 2017, the difference in gains attained between T1 and T2 had increased, and gains can be observed beyond the lower secondary level, driven by those ages 1-5 at the start of the program. As before, for interpretation we need to account for the fact that these are results for children with siblings in our 2 main cohorts. For easier comparison, we therefore show results for the early childhood cohort restricting to households where there was another sibling in the main sample. These results are shown in the top panel (note that in 2003 this cohort was too young for schooling data to be collected so that they could not be included in the medium-term analysis in Table C4 and C5), confirming that the overall gains in educational attainment of the main sample are somewhat larger than that of their siblings in the intermediate age group. Together with the results by age-cohort discussed earlier, the findings show that children generally benefitted from earlier exposure anytime between conception and age 5, consistent with the age groups typically considered in early childhood interventions. The pattern of the results further suggests that children who started receiving transfers earlier benefited more, though differences are not significant, possibly due to lack of power.

Appendix D: Data collection and tracking

Data collection

Fieldwork was conducted in two phases. The first phase took place between May and August 2017 and the second one between September and December 2017. During the first phase, the enumerators visited all the villages in the sample. In coordination with the program's voluntary village-level leaders (*vocales*), a community meeting was convened. All the main beneficiaries (*titulares*) from the households to which the individuals in the tracking sample belonged were invited. In most cases, they were the mothers of the individuals in the tracking sample. The meeting had three objectives. First, to explain the purpose of the survey and gain families' trust and willingness to participate in the study. Second, to identify which mothers of the individuals in the sample still lived in the village. Third, to schedule appointments with those living in the village to conduct the survey. And fourth, to gather any information on families from the sample who no longer lived in the village.

In the cases where the individuals in the tracking sample still lived in the village, the enumerator visited their homes and conducted a survey interview with them. The survey administration took approximately 20-25 minutes. It included information on their educational attainment, marital status, ages when they formed a family and had children, number of children, occupation and labor income, migration history, and recall information on past Progresa benefits. For those with a partner, it also inquired about their partner's educational attainment, age, place of origin, and employment. For those in the early childhood cohort, many of whom were still studying, the questionnaire included questions on their educational and income expectations. For those in the school cohort, who were more likely to have finished all their education, entered the labor market, and formed their own household, the survey also asked information on asset ownership. To allow for survey administration by phone, questionnaires were purposely kept short, and notably could not include anthropometric measures, cognitive or achievement (learning) tests.

In the cases where the individuals in the tracking sample were no longer in the village, we looked for their mothers or, in absence of the mother, another member of their baseline (1997) household. If they still lived in the village, the enumerator visited their home to conduct an interview with them as proxy informants. The questionnaire for these proxy interviews was shorter (approximately 10-15 minutes long) and less detailed than the one applied to the individuals directly. It included questions on the individual's schooling, employment, marital status, number of children, and in cases where the individual in the tracking sample had formed a family, on the schooling and employment of his/her partner. In addition, enumerators gathered all available information on the current location of the individual, including how to contact him/her by phone, residence and work addresses, and social media contacts. Finally, at least one respondent in every household was asked about other household members from their 1997 household, including whether they had migrated (and where), their age, schooling, and labor market participation.

In cases that individuals in our two cohorts were not found in their village of origin, the tracking protocol encouraged enumerators to call the individuals while still in the company of their mother (or relative) and asked the mother herself to introduce the survey so that it would gain the support of the individual to participate in it later. Experience during piloting confirmed afterwards in the study, showed that the most successful tracking occurred when a follow-up phone interview could be conducted immediately or shortly afterwards the visit to the home (e.g., later in the same day) with the individual by phone, rather than by trying to trace them for a face-to-face interview. When the individual in the tracking sample was contacted by phone, the complete questionnaire described above was administered. Tracking of migrants that were in the US occurred by phone. Some international migrants were interviewed in person or by phone once they returned to Mexico.

When immediate phone contact with respondents was not feasible (because the household of origin could not be found, had no phone number, or the phone number they provided did not work), enumerators searched for other knowledgeable people in the village (family members, friends, leaders, etc.) to cross-validate destination and contact information. Prior to starting any of the interviews or attempts of obtaining contact information, a 3-minute video message, recorded by the program's national coordinator, was shown to explain the importance of the survey. The videos were shown on tablets before the face-to-face interviews, or sent through WhatsApp, email, or Facebook prior to phone surveys.

The search and interview efforts during the second phase of the data collection were organized based on the information gathered during the first phase. Some field teams went back to the villages in our sample while others looked for the respondents in urban areas and reported destinations across Mexico. Prior to the field visits, personalized invitation letters issued by the program to participate in the interview were sent to the villages of origin. These letters were also shown (together with the video message) when contacting people at their destination and were further used to show to employers (or others), in cases where their authorization had to be obtained to conduct and interview.

The second phase involved different activities, carried out simultaneously. First, there was an active effort to return to 150 of the original villages to a) recover information on individuals for whom no or incomplete contact information was found during the first phase, including through wider family and community networks and acquaintances; and b) interview respondents living in the village who were absent during the first phase. Second, we attempted to find all the individuals who had moved out of their villages and for whom we had addresses of their new residences anywhere in Mexico. This implied tracking individuals anywhere in the 7 original states, as well as individuals who had moved to 23 other states across Mexico. In cases where the search was successful, we collected in-person interviews to complement (and validate the quality of) the information provided previously by the proxy informants. Third, when new phone contact information was obtained, additional phone interviews were conducted. Fourth, additional efforts were made to contact individuals through phone numbers obtained in the first phase, including international migrants. Fifth, for a subsample of

99 individuals who had been interviewed by phone and who were willing to answer an in-person survey, the team visited them in their homes and conducted a face-to-face interview with the aim of validating the quality of the survey data collected via phone interviews. Sixth, every individual that was traced in person or by phone, was asked about the location and contact information of other missing respondents of his/her original household, and when feasible, about other missing respondents of their origin village.

The reliance on the program's structure and social capital implies that the likelihood of success in finding an individual could be related to the family´s status in Prospera. However, as discussed in the main text and shown in Table 1, there is no difference between T1 and T2 in the probability of success of tracking. Individuals from households that moved out of the original localities long time ago are less likely to be found (as obtaining contact information about them was more challenging) and also less likely to still be active in the program, but this is the true for T1 and T2. Attrition is also not correlated to missing baseline information (for 5 percent of individuals have the 1997 ENCASEH is missing), and non-attrited observations show balance on baseline observables (Table D1). [6]

Self-reported vs. proxy data

One key question related to the data used for the analysis is, to what extent the proxy information reported by the mother or closest relative (i.e., a member of the original household of the individual in the tracking sample) is accurate. To validate the quality of the proxy information collected we compare proxy information to self-reported information for individuals for whom both was collected. This was the case when at the first tracking attempt, a survey was completed with the mother and then, once the individual was found (either in the first or the second phase), a new survey was conducted in which the beneficiary answered directly. There are 2463 individual for whom we have both a proxy and a direct survey.

Table D2 shows the comparison of answers on education, household formation, and occupation as these were asked to both types of respondents in similar ways. A similar comparison cannot be done for most information on geographic mobility because it was not collected from the proxy informant in cases immediate contact by phone with the main beneficiary was established after the proxy survey. And when such phone contact was not possible, information on the migration destination was collected in a separate module building on extensive triangulation and cross-validation during the first phase of the survey, so that we do not have a "clean" measure

[6] Broadly in line with expectations, of the 19 baseline variables in Table D1 there are 2 significant (at the 10%) differences between T1 and T2 for the school cohort, 3 for the full early childhood cohort, and 1 (at the 5%) for the early childhood cohort excluding 1999. We have also tested robustness of all results using post double selection lasso estimates (Belloni et al 2014) to account for accidental imbalances (results available from authors).

of the proxy report for the location information. The proxy and self-reported survey did however ask one question on migration history (whether ever migrated to the US) so this variable can directly be compared.

Before comparing answers of the direct survey and the proxy respondent, we need to account for the fact that there are some variables for which a relatively large number of proxy respondents indicated they could not answer. Table D2 therefore shows in the first column the percentage of individuals for whom either the self-reported or the proxy response to a particular question was missing (in the vast majority of cases this will be due to non-response on that variable in the proxy survey). Non-response is a minor concern for most of the variables, with the notable exception of the information on the beneficiary's partner. A relatively large share of proxy respondents indicated they did not know the age or education of their children's partner, which was particularly the case when children had migrated. Given the possible selection concerns that this high level of non-response among the proxy reports could create, we interpret the results on beneficiaries' partners more cautiously and only report them for completeness in Appendix E. Non-response is less of a concern for the other variables, 16 percent for questions asking for the age at marriage and age when first child was born, and less than 2 percent for all other variables.

The second column of Table D2 compares the self-reported to the proxy information. For continuous variables we present Pearson correlations (in the top panel), for binary variables we report the percentage of responses that was identical across the two data sources (in the bottom panel). Correlations are above 0.85 for grades attained, number of children and age of the partner, and between 0.076 and 0.78 for the age at which the individual had his or her first child, age at marriage and grades attained by the partner. Among all of the dummy variables, the percentage of coincidence in the values reported is close to or above 90 percent for almost all variables, and slightly lower (83 percent) for whether the target individual is the household head or spouse of the head in their current household. Overall, these comparisons show that the coincidence among the sources is high (and indeed much higher than the 0.70 threshold often used in psychometrics for test-retest statistics), confirming the reliability of the proxy report in this sample.

In-person vs. phone interviews

Another related data quality concern has to do with the quality of information collected by phone rather than through a face-to-face interview. In this case, we refer to surveys that were answered by the same informant, but through a phone interview. To document how reliable this data was, during the second stage of data collection, we selected a random sub-sample of 300 individuals interviewed by phone and attempted to find them for a face-to-face interview. Such tracking was not only complicated and costly, but often security concerns made individuals reluctant to provide detailed address information that allowed to locate individuals in person, or where otherwise unwilling to participate in the survey a second time. As a result, we only have 99 cases to compare information collected through phone versus in-person interviews. With the caveat of the comparison between the face-to-face and phone survey hence being done on a selected sample only, it is still presented here, as it also it provides a benchmark for the comparison of the proxy with the direct answers.

The last column of Table D2 compares the information reported through in-person versus phone interviews for the subsample of respondents for whom it was possible to apply the survey twice. The correlation coefficients and percentages of coinciding values are also high and broadly of similar magnitudes to the ones presented earlier. Overall, results confirm the reliability of most of the information obtained through the different informants (individual or proxy) and data collection channels (in person or phone interviews).

Appendix E: Additional results

Early childhood results by birthyear

In an attempt to explore whether we can discern particularly sensitive periods for the effect of differential exposure to Progresa within the first 1000-day window, we estimated a specification with separate treatment effects for individuals in the early childhood cohort born in 1997, 1998 and 1999, and tested the differences between them. Children born in 1997 received benefits starting age 5-16 months if they were in T1, but only at age 22-34 months if they were in T2. Children born in 1998 received benefits starting age -8 to 4 months if in T1, while they already were age 10-22 months when treatment started in T2. And children born in 1999 were exposed in utero if they were part of a T1 household, while they were only exposed after birth if they were in T2.

Table E1 shows that there are no significant differences across birthyears in either educational gains or income expectations. This suggests that being exposed to Progresa earlier in life during the first 1000-day window led to longer-term gains, without clear differences within that window. That said, the comparison between birthyear cohorts needs to be interpreted with caution, as many more individuals born in 1999 were still in school in 2017, compared to those born in 1997. Moreover, we cannot exclude that differences between birthyears are hard to pick up in our design due to lack of statistical power (which was the rationale for pulling the three birth years in the main analysis). It is therefore interesting to note that the point estimates of the differential treatment effect for having completed lower secondary school (the highest level that all three birthyears could have finished given their age) are larger for those born in 1998 or 1999 (0.061 and 0.069 compared to 0.038), i.e., those for whom differential exposure included the in-utero period. Results on income expectations point to possibly larger gains for those born in 1998.

Other outcomes for the early childhood cohort

For completeness, we present the results for the rest of the outcomes collected for the early childhood cohort. These are harder to interpret, as this cohort was too young in 2017 to have reached final outcomes on economic activities, mobility, fertility, and marriage. Nevertheless, we note that for total income there seems to be a positive differential (significant when trimming outliers), which is not related to US income (Table E2). As for the school cohort, we find no differences in occupations and job characteristics (available only for the 55% of this cohort that is already economically active).

We find no significant differences in marriage status or having started childbearing, with the mean in the control at 28 and 21% (Table E3). This is in line with the results of the older cohort, for whom differences in fertility and marriage occurred after age 20. There are no significant differences in health-related behavior (smoking and drinking), though these outcomes may suffer from social desirability bias.

Finally, given the low proportion of married individuals, and concerns with measurement error, differences in partner characteristics are at best suggestive. Possibly interesting to note is that the early childhood cohort - particularly women- when married, are less likely to have lower education than their partners and are more likely to have a partner from a different locality or municipality (11 p.p. and 8 p.p., respectively). Men from T1, on the other hand, are more likely to have a partner with completed lower secondary school (13 p.p.) and to have a partner from a different country (3 p.p.), from a very low mean. There are also some suggestive differences in the types of occupations of their partners.

Other outcomes for the school cohort

Table E4 and E5 report some additional outcome variables on occupation and job characteristics, already discussed in the main text. Table E6 reports differences in partner characteristics. While the share of people in a relationship is relatively high and balanced in the school cohort, results in Table C2 indicate the partner characteristics suffer from more selection (more missing values) and measurement error. Table E6 is therefore mostly reported for completeness. Overall, the table shows few significant differences between T1 and T2. At the bottom of the table, we also include results on smoking and drinking behavior, but caveat these are likely affected by social desirability bias.

Table A1: Compliance and program participation of school cohort

	All		**Women**		**Men**		**P-value**
	Obs	coef.	Obs	coef.	Obs	coef.	Women=Men
	(1)		(2)		(3)		(4)
Administrative data							
No transfer data for household in administrative disbursement datasets							
T1	2153	-0.02	1097	-0.004	1056	-0.037	0.138
s.e.		(0.012)*		(0.014)		(0.018)**	
mean T2		0.069		0.052		0.087	
Actual Cumulative Transfer between 1998 and 2003							
T1	1908	1130.853	984	1108.607	924	1143.684	0.972
s.e.		(116.838)***		(130.503)***		(161.753)***	
mean T2		2714.155		2772.816		2650.769	
Original household still active in Prospera registry in 2017							
T1	2153	0	1097	-0.026	1056	0.028	0.239
s.e.		(0.028)		(0.035)		(0.036)	
mean T2		0.594		0.637		0.548	
Individual listed as active in 2017 Prospera registry							
T1	2153	-0.019	1097	-0.022	1056	-0.019	0.995
s.e.		(0.024)		(0.03)		(0.03)	
mean T2		0.23		0.215		0.246	
Listed as titular or household head in 2017 Prospera registry							
T1	2153	-0.002	1097	-0.003	1056	-0.002	0.908
s.e.		(0.007)		(0.011)		(0.008)	
mean T2		0.029		0.038		0.02	
Self-reported data							
Report growing up in household receiving Progresa/Prospera							
T1	1884	0.141	1004	0.155	880	0.124	0.432
s.e.		(0.023)***		(0.031)***		(0.032)***	
mean T2		0.743		0.734		0.755	
Report receiving Progresa/Prospera school transfers							
T1	1690	0.188	914	0.231	776	0.138	0.055*
s.e.		(0.028)***		(0.035)***		(0.039)***	
mean T2		0.625		0.597		0.659	
Number of years of Progresa/Prospera school transfers, self-report							
T1	1690	0.987	914	1.148	776	0.78	0.219
s.e.		(0.202)***		(0.255)***		(0.268)***	
mean T2		3.074		3.003		3.162	
Age at which received first Progresa/Prospera school transfer, self-report							
T1	1305	-0.267	699	-0.326	606	-0.192	0.528
s.e.		(0.175)		(0.241)		(0.235)	
mean T2		10.752		10.973		10.512	
Age at which received last Progresa/Prospera school transfer, self-report							
T1	1345	-0.241	720	-0.119	625	-0.362	0.451
s.e.		(0.169)		(0.216)		(0.23)	
mean T2		15.153		15.155		15.151	
Reports receiving Progresa/Prospera school transfer when 18 years old							
T1	1345	-0.05	720	-0.052	625	-0.046	0.892
s.e.		(0.026)*		(0.037)		(0.035)	
mean T2		0.216		0.235		0.195	
Reports living in hh that receives benefits Prospera in 2017							
T1	1869	-0.002	1000	0.009	869	-0.013	0.575
s.e.		(0.029)		(0.037)		(0.034)	
mean T2		0.36		0.381		0.334	
Reports living in original hh where titular receives benefits Prospera in 2017							
T1	1869	0.007	1000	0.004	869	0.009	0.946
s.e.		(0.022)		(0.028)		(0.028)	
mean T2		0.179		0.169		0.191	

Notes: ITT estimates with region and village population size fixed effects. Col. 1 shows results for the full schooling cohort, Col. 2 for female and Col. 3 for males. Col. 4 shows the p-values for tests that the effects are identical between males and females. Each column shows the number of observations and the estimated differential treatment effect in the first row, standard errors are in the second cell row parentheses, clustered at locality level, with ***, **, * indicating significance at 1, 5 and 10%; and mean in T2 localities in the third row.

Table A2: Compliance and program participation of early childhood cohort

	All		Women		Men		P-value	Born in 1997 and 1998	
	Obs	coef.	Obs	coef.	Obs	coef.	Women=Men	Obs	coef.
	(1)		(2)		(3)		(4)	(5)	
Administrative data									
No transfer data for household in administrative disbursement datasets									
T1	4137	-0.02	2083	-0.02	2054	-0.02	0.864	3036	-0.02
s.e.		(0.012)*		(0.016)		(0.015)			(0.015)
mean T2		0.106		0.111		0.101			0.128
Actual Cumulative Transfer between 1998 and 2003									
T1	3447	567.999	1721	665.32	1726	475.413	0.113	2448	612.156
s.e.		(81.749)***		(91.213)***		(112.134)***			(84.286)***
mean T2		1909.638		1806.654		2017.188			1897.798
Original hh still active in Prospera registry in 2017									
T1	4137	-0.002	2083	-0.021	2054	0.017	0.379	3036	-0.015
s.e.		(0.023)		(0.027)		(0.027)			(0.025)
mean T2		0.61		0.615		0.606			0.59
Individual listed as active in 2017 Prospera registry (not including those born in 98/99)									
T1	1363	-0.002	697	0.015	666	-0.024	0.552		
s.e.		(0.03)		(0.041)		(0.042)			
mean T2		0.456		0.419		0.494			
Listed as titular or hh head in 2017 Prospera registry (not including those born in 98/99)									
T1	1363	-0.004	697	-0.006	666	-0.002	0.613		
s.e.		(0.004)		(0.006)		(0.004)			
mean T2		0.006		0.007		0.004			
Self-reported data									
Report growing up in household receiving Progresa/Prospera									
T1	3831	0.034	1968	0.036	1863	0.032	0.79	2782	0.039
s.e.		(0.012)***		(0.016)**		(0.015)**			(0.013)***
mean T2		0.879		0.877		0.881			0.873
Report receiving Progresa/Prospera school transfers									
T1	3542	0.029	1829	0.025	1713	0.035	0.782	2565	0.03
s.e.		(0.015)*		(0.019)		(0.02)*			(0.017)*
mean T2		0.856		0.858		0.854			0.854
Number of years of Progresa/Prospera school transfers, self-report									
T1	3542	0.547	1829	0.655	1713	0.442	0.369	2565	0.628
s.e.		(0.188)***		(0.215)***		(0.233)*			(0.202)***
mean T2		5.978		5.934		6.028			5.872
Age at which received first Progresa/Prospera school transfer, self-report									
T1	3156	-0.087	1631	-0.241	1525	0.072	0.069*	2289	-0.146
s.e.		(0.084)		(0.108)**		(0.12)			(0.102)
mean T2		8.615		8.675		8.546			8.701
Age at which received last Progresa/Prospera school transfer, self-report									
T1	3237	0.17	1672	0.262	1565	0.066	0.213	2351	0.165
s.e.		(0.155)		(0.179)		(0.179)			(0.168)
mean T2		15.358		15.318		15.404			15.36
Reports receiving Progresa/Prospera school transfer when 18 years old									
T1	3237	0.028	1672	0.044	1565	0.012	0.233	2351	0.032
s.e.		(0.027)		(0.031)		(0.032)			(0.03)
mean T2		0.283		0.272		0.295			0.294
Reports living in hh that receives benefits Prospera in 2017									
T1	3783	-0.004	1953	-0.014	1830	0.003	0.668	2746	0.005
s.e.		(0.022)		(0.026)		(0.028)			(0.024)
mean T2		0.615		0.595		0.638			0.576
Reports living in original hh where titular receives benefits Prospera in 2017									
T1	3783	-0.005	1953	-0.005	1830	-0.013	0.734	2746	0.005
s.e.		(0.022)		(0.026)		(0.028)			(0.025)
mean T2		0.521		0.467		0.582			0.479

Notes: ITT estimates with region and village population size fixed effects. Col. 1 shows results for the full early childhood cohort, Col. 2 for female and Col. 3 for males. Col. 4 shows the p-values for tests that the effects are identical between males and females. Each column shows the number of observations and the estimated differential treatment effect in the first row, standard errors are in the second cell row parentheses, clustered at locality level, with ***, **, * indicating significance at 1, 5 and 10%; and mean in T2 localities in the third row.

Table C1: 2003 differential impacts on main cohorts

	Currently in school	Literate (reads and writes)	Grades attained	Level completed: Primary	Level completed: Lower secondary	Level completed: > Lower secondary	Currently in school
Born in 1997			**9-15 year olds with 5 grades attained & enrolled in 1997**				
ENCEL sample							
T1	-0.0101	0.0581**	0.397***	-0.00203	0.108***	0.00954	-0.0282
s.e.	(0.015)	(0.029)	(0.13)	(0.0088)	(0.034)	(0.025)	(0.034)
mean T2	0.955	0.226	7.81	0.956	0.501	0.183	0.321
Obs.	1014	1182	1946	1946	1946	1946	1291
2017 tracking sample							
T1	0.00165	0.0632**	0.395***	0.00170	0.108***	0.0122	-0.0193
s.e.	(0.013)	(0.030)	(0.14)	(0.0096)	(0.034)	(0.026)	(0.035)
mean T2	0.959	0.228	7.831	0.948	0.506	0.186	0.319
Obs.	887	1037	1891	1891	1891	1891	1249

Note: Estimates based on 2003 ENCEL data. ITT estimates with region and village population size fixed effects, standard errors are in the second cell row parentheses, clustered at locality level, with ***, **, * indicating significance at 1, 5 and 10%

Table C2: 2003 differential impacts on other cohorts of school ages

	Grades attained	Level completed: Primary	Level completed: Lower secondary	Level completed: > Lower secondary	Still studying
9-15 year olds with 4 or less grades attained & enrolled in 1997					
T1	0.270**	0.0544***	0.0433**	0.00404	-0.0102
	(0.11)	(0.015)	(0.022)	(0.0089)	(0.027)
mean T2	7.059	0.831	0.351	0.0644	0.439
Observations	5995	5995	6000	5995	4908
9-15 year olds with 6 or more grades attained & enrolled in 1997					
T1	0.210*	n.a.	0.0327*	-0.00891	-0.0406
	(0.12)		(0.020)	(0.031)	(0.032)
mean T2	9.117		0.804	0.277	0.249
Observations	2498		2503	2498	1333
9-15 year olds not enrolled in 1997					
T1	0.404***	0.0752***	0.0472**	0.00920	0.0168
	(0.14)	(0.027)	(0.019)	(0.0062)	(0.015)
mean T2	5.419	0.666	0.142	0.0218	0.0765
Observations	2951	2951	2956	2951	1658
16-17 year olds in 1997					
T1	0.193	0.0196	0.0244	-0.00199	-0.00710
	(0.19)	(0.025)	(0.029)	(0.018)	(0.017)
mean T2	6.623	0.768	0.338	0.104	0.0722
Observations	2320	2320	2323	2320	946

Note: Estimates based on 2003 ENCEL data. ITT estimates with region and village population size fixed effects, standard errors are in the second cell row parentheses, clustered at locality level, with ***, **, * indicating significance at 1, 5 and 10%

Table C3: 2003 differential impacts on siblings of school ages

	Grades attained	Level completed			Still studying
		Primary	Lower secondary	> Lower secondary	
9-15 year olds with 4 or less grades attained & enrolled in 1997					
T1	0.417***	0.0848***	0.0638**	0.000819	-0.0150
	(0.11)	(0.016)	(0.027)	(0.011)	(0.033)
mean T2	6.899	0.805	0.318	0.0551	0.449
Observations	2536	2536	2537	2536	2112
9-15 year olds with 6 or more grades attained & enrolled in 1997					
T1	0.218	0.0150	0.0350	0.0262	0.00278
	(0.18)	(0.011)	(0.032)	(0.036)	(0.044)
mean T2	9.046	0.966	0.783	0.239	0.227
Observations	981	981	982	981	486
9-15 year olds not enrolled in 1997					
T1	0.450***	0.105***	0.0250	0.0111	0.0178
	(0.17)	(0.032)	(0.027)	(0.0100)	(0.025)
mean T2	5.54	0.672	0.162	0.0204	0.0798
Observations	1150	1150	1154	1150	607
16-17 year olds in 1997					
T1	0.0938	-0.00335	0.0372	-0.0176	-0.0236
	(0.23)	(0.032)	(0.037)	(0.023)	(0.029)
mean T2	6.77	0.796	0.331	0.116	0.0872
Observations	964	964	965	964	359

Note: Estimates based on 2003 ENCEL data. ITT estimates with region and village population size fixed effects, standard errors are in the second cell row parentheses, clustered at locality level, with ***, **, * indicating significance at 1, 5 and 10%

Table C4 : 2017 differential impacts on siblings of school ages

	Grades attained	Level completed			Still studying	In US in 2017	Living in locality of origin
		Primary	Lower secondary	> Lower secondary			
9-15 year olds with 4 or less grades attained & enrolled in 1997							
T1	0.420**	0.0798***	0.0585*	-0.00453	-0.000370	0.00682	-0.00645
	(0.19)	(0.018)	(0.032)	(0.0095)	(0.0037)	(0.018)	(0.028)
mean T2	7.789	0.832	0.513	0.0418	0.008	0.0955	0.446
Observations	2454	2454	2454	2454	2424	2566	2572
9-15 year olds with 6 or more grades attained & enrolled in 1997							
T1	0.184	0.00320	0.0285	-0.00162	0.00116	0.00943	0.0186
	(0.23)	(0.012)	(0.034)	(0.020)	(0.0070)	(0.028)	(0.036)
mean T2	9.239	0.971	0.763	0.0737	0.01	0.136	0.361
Observations	959	959	959	959	939	1008	1012
9-15 year olds not enrolled in 1997							
T1	0.332	0.0779***	0.0162	-0.00616	-0.00423	0.00625	-0.0397
	(0.21)	(0.030)	(0.032)	(0.0055)	(0.0037)	(0.022)	(0.034)
mean T2	5.819	0.71	0.211	0.0117	0.006	0.0991	0.5
Observations	1183	1183	1183	1183	1183	1280	1284
16-17 year olds in 1997							
T1	0.315	0.0457	0.0228	0.0119	0.00132	-0.00442	-0.0316
	(0.26)	(0.032)	(0.037)	(0.011)	(0.0014)	(0.022)	(0.036)
mean T2	6.749	0.785	0.364	0.0205	0	0.0982	0.486
Observations	972	972	972	972	978	1048	1051

Note: Estimates based on 2017 tracking data. ITT estimates with region and village population size fixed effects, standard errors are in the second cell row parentheses, clustered at locality level, with ***, **, * indicating significance at 1, 5 and 10%

Table C5: 2003 differential impacts on cohorts of intermediate ages

	Grades attained	Currently in school	Literate (reads and writes)
All children 1-8 year old in 1997			
T1	0.0881*	-0.00144	0.00558
	(0.050)	(0.0087)	(0.0083)
mean T2	3.807	0.926	0.85
Observations	18301	17819	19019
Children 1-2 year old in 1997			
T1	0.0288	-0.0102	0.00473
	(0.033)	(0.0069)	(0.019)
mean T2	1.634	0.98	0.716
Observations	3676	3596	3882
Children 3-5 year old in 1997			
T1	0.0791*	-0.00381	-0.000751
	(0.043)	(0.0052)	(0.0085)
mean T2	3.343	0.98	0.944
Observations	6595	6361	6709
Children 6-8 year old in 1997			
T1	0.191**	0.00592	0.00565
	(0.076)	(0.018)	(0.0065)
mean T2	5.682	0.836	0.97
Observations	6805	6507	6868

Note: ITT estimates with region and village population size fixed effects, standard errors are in the second cell row parentheses, clustered at locality level, with ***, **, * indicating significance at 1, 5 and 10%

Table C6: 2003 differential impacts on siblings of intermediate ages

	Grades attained	Currently in school	Literate (reads and writes)
All children 1-8 year old in 1997			
T1	0.136**	-0.00845	0.00872
	(0.058)	(0.0092)	(0.0087)
mean T2	3.841	0.935	0.891
Observations	8268	7966	8473
Children 1-2 year old in 1997			
T1	0.0865**	0.00221	0.0440*
	(0.042)	(0.0075)	(0.026)
mean T2	1.611	0.98	0.688
Observations	1769	1720	1865
Children 3-5 year old in 1997			
T1	0.111**	-0.00644	-0.00522
	(0.051)	(0.0050)	(0.0091)
mean T2	3.328	0.988	0.951
Observations	3205	3089	3262
Children 6-8 year old in 1997			
T1	0.251***	-0.0144	0.000759
	(0.084)	(0.021)	(0.0062)
mean T2	5.664	0.853	0.978
Observations	3166	3027	3194

Note: ITT estimates with region and village population size fixed effects, standard errors are in the second cell row parentheses, clustered at locality level, with ***, **, * indicating significance at 1, 5 and 10%

Table C7: 2017 differential impacts on siblings of intermediate ages

	Grades attained	Level completed			
		Primary	Lower secondary	Upper secondary	Some university
Young cohort (born between 1997 and 1999) -with other sibling in tracking survey					
T1	0.507**	0.0180	0.104***	0.0384	0.0217**
	(0.21)	(0.015)	(0.034)	(0.030)	(0.011)
mean T2	8.981	0.934	0.685	0.245	0.0237
Observations	1514	1514	1514	1514	1514
<u>All</u> children 1-8 year old in 1997					
T1	0.362*	0.0236*	0.0391	0.0421	0.0165**
	(0.18)	(0.013)	(0.027)	(0.027)	(0.0080)
mean T2	8.781	0.917	0.655	0.24	0.0538
# observations	8235	8235	8235	8235	8235
Children 1-2 year old in 1997					
T1	0.446**	0.0179	0.0466	0.0755**	0.00652
	(0.21)	(0.017)	(0.033)	(0.031)	(0.013)
mean T2	8.964	0.927	0.682	0.265	0.0633
Observations	1844	1844	1844	1844	1844
Children 3-5 year old in 1997					
T1	0.341*	0.0158	0.0350	0.0384	0.0223**
	(0.20)	(0.013)	(0.029)	(0.030)	(0.010)
mean T2	8.852	0.923	0.664	0.246	0.0549
Observations	3258	3258	3258	3258	3258
Children 6-8 year old in 1997					
T1	0.330	0.0352**	0.0381	0.0247	0.0154
	(0.20)	(0.017)	(0.030)	(0.028)	(0.0094)
mean T2	8.603	0.906	0.63	0.219	0.0472
Observations	3133	3133	3133	3133	3133

Note: ITT estimates with region and village population size fixed effects, standard errors are in the second cell row parentheses, clustered at locality level, with ***, **, * indicating significance at 1, 5 and 10%

Table D1: Baseline variables and balance for different cohorts and samples

	Early childhood cohort								School cohort			
	Born 1997-1999				Born 1997-1998							
	All tracked		Self-report only		All tracked		Self-report only		All tracked		Self-report only	
	Obs	Coef.	Obs	Coef.	Obs	Coef.	Obs	Coef.	Obs	Coef.	Obs	Coef.
Household size												
T1	3880	-0.138	3595	-0.138	2796	-0.077	2562	-0.071	2088	-0.128	1826	-0.113
s.e.		(0.134)		(0.131)		(0.145)		(0.143)		(0.113)		(0.12)
mean T2		6.46		6.411		6.516		6.444		7.48		7.458
Baseline proxy means poverty index												
T1	3872	1.946	3587	0.669	2790	3.454	2556	2.141	2080	2.748	1819	4.174
s.e.		(5.587)		(5.535)		(5.869)		(5.876)		(6.429)		(6.269)
mean T2		625.839		626.996		623.541		624.535		623.689		623.076
Household receives food subsidy												
T1	3880	0.032	3595	0.039	2796	0.028	2562	0.033	2088	0.027	1826	0.028
s.e.		(0.034)		(0.034)		(0.034)		(0.034)		(0.041)		(0.042)
mean T2		0.33		0.326		0.335		0.331		0.448		0.445
Household accesss IMSS health services												
T1	3880	0.04	3595	0.037	2796	0.056	2562	0.052	2088	0.007	1826	0
s.e.		(0.039)		(0.04)		(0.04)		(0.041)		(0.043)		(0.043)
mean T2		0.304		0.308		0.297		0.302		0.38		0.397
House with cardboard roof												
T1	3880	-0.023	3595	-0.021	2796	-0.03	2562	-0.028	2088	-0.017	1826	-0.027
s.e.		(0.029)		(0.028)		(0.03)		(0.03)		(0.028)		(0.029)
mean T2		0.256		0.256		0.267		0.266		0.203		0.207
House with access to running water												
T1	3880	0.064	3595	0.064	2796	0.068	2562	0.063	2088	0.082	1826	0.081
s.e.		(0.041)		(0.041)		(0.042)		(0.043)		(0.046)*		(0.046)*
mean T2		0.249		0.248		0.256		0.257		0.278		0.277
House with electricity												
T1	3880	-0.007	3595	-0.008	2796	-0.01	2562	-0.008	2088	-0.043	1826	-0.032
s.e.		(0.047)		(0.047)		(0.048)		(0.048)		(0.044)		(0.045)
mean T2		0.585		0.584		0.591		0.589		0.686		0.678
Household owns land												
T1	3880	0.026	3595	0.025	2796	0.047	2562	0.048	2088	0.004	1826	-0.004
s.e.		(0.035)		(0.036)		(0.037)		(0.037)		(0.033)		(0.034)
mean T2		0.54		0.537		0.519		0.513		0.666		0.668
Household owns a working animal												
T1	3880	0.031	3595	0.023	2796	0.037	2562	0.03	2088	0.025	1826	0.015
s.e.		(0.026)		(0.027)		(0.027)		(0.028)		(0.032)		(0.034)
mean T2		0.311		0.314		0.304		0.304		0.395		0.394

Household owns cattle												
T1	3880	0.036	3595	0.038	2796	0.045	2562	0.049	2088	0.007	1826	0.011
s.e.		(0.019)*		(0.019)**		(0.02)**		(0.02)**		(0.023)		(0.024)
mean T2		0.129		0.127		0.117		0.113		0.185		0.178
Age household head												
T1	3880	0.066	3595	-0.051	2796	0.467	2562	0.411	2088	-0.902	1826	-0.503
s.e.		(0.589)		(0.596)		(0.6)		(0.606)		(0.534)*		(0.593)
mean T2		37.383		37.21		36.962		36.734		44.7		44.455
Female household head												
T1	3877	0.003	3592	0.005	2793	0.008	2559	0.011	2087	0	1825	0.003
s.e.		(0.009)		(0.008)		(0.009)		(0.009)		(0.013)		(0.014)
mean T2		0.051		0.048		0.045		0.042		0.069		0.068
Household head completed primary												
T1	3859	0.044	3574	0.047	2784	0.028	2550	0.029	2076	0.037	1817	0.03
s.e.		(0.024)*		(0.024)**		(0.026)		(0.026)		(0.022)		(0.023)
mean T2		0.25		0.255		0.266		0.27		0.145		0.15
Household head completed lower secondary												
T1	3859	0.017	3574	0.018	2784	0.015	2550	0.015	2076	0	1817	0.001
s.e.		(0.009)*		(0.009)*		(0.01)		(0.011)		(0.008)		(0.009)
mean T2		0.037		0.039		0.042		0.044		0.024		0.025
Household head speaks an indigenous language												
T1	3875	-0.018	3591	-0.012	2793	-0.007	2560	0.003	2087	-0.07	1825	-0.064
s.e.		(0.058)		(0.058)		(0.057)		(0.057)		(0.054)		(0.055)
mean T2		0.453		0.45		0.44		0.434		0.45		0.457
Household head is agricultural laborer												
T1	3880	-0.035	3595	-0.028	2796	-0.036	2562	-0.03	2088	-0.039	1826	-0.037
s.e.		(0.03)		(0.03)		(0.03)		(0.03)		(0.033)		(0.034)
mean T2		0.634		0.631		0.635		0.633		0.586		0.586
Household head has non-agricultural employment												
T1	3880	-0.028	3595	-0.028	2796	-0.031	2562	-0.031	2088	-0.016	1826	-0.017
s.e.		(0.019)		(0.019)		(0.02)		(0.021)		(0.017)		(0.019)
mean T2		0.096		0.097		0.1		0.101		0.08		0.081
Household head started working by age 14												
T1	3880	-0.023	3595	-0.027	2796	-0.016	2562	-0.018	2088	0.025	1826	0.025
s.e.		(0.023)		(0.024)		(0.025)		(0.025)		(0.026)		(0.028)
mean T2		0.565		0.563		0.562		0.559		0.575		0.577
Household head migrated for work in last 12 months												
T1	3880	0.015	3595	0.023	2796	0.015	2562	0.024	2088	-0.007	1826	-0.016
s.e.		(0.018)		(0.018)		(0.019)		(0.019)		(0.017)		(0.017)
mean T2		0.098		0.094		0.098		0.091		0.091		0.093

Notes: ITT estimates with region and village population size fixed effects. Col. 1 and 2 shows results for the full early childhood cohort, Col. 3 and 4 shows results restricting to those born in 97 and 98, Col. 5 and 6 shows results for the schooling cohort.Uneven columns baesd on sample for whom information was collected at follow-up through self-report or proxy reporting. Even columns restrict sample to those for whom data was collected at follow-up through self-report only. All columns exclude observations with missing baseline data (5% of total). Each column shows the number of observations and the estimated differential treatment effect in the first row, standard errors are in the second cell row parentheses, clustered at locality level, with ***, **, * indicating significance at 1, 5 and 10%; and mean in T2 localities in the third row. Baseline data is missing for 5% of individuals.

Table D2: Comparison of different survey methods

	% with missing answer in proxy interview	Correlation direct & proxy answers	Correlation face-to-face & phone answers
Continuous and count variables			
Grades attained (years of education)	1	0.87	0.89
Number of children	1	0.89	0.98
Age when had first child	16	0.78	0.71
Grades attained by partner	38	0.78	0.79
		Coincidence direct & proxy answers	Coincidence face-to-face & phone answers
Binary variables			
Completed primary school	1	0.97	0.99
Completed lower secondary school	1	0.92	0.94
Has income in US$	1	0.95	0.91
Completed some tertiary education	1	0.97	0.96
Still studying	1	0.97	0.92
Ever married or in civil union	0	0.96	0.96
Has any children	1	0.96	0.97
Is household head or spouse of head	1	0.83	0.83
Ever migrated to USA	0	0.95	0.96
Participated in labor market in last 12 months	2	0.88	0.80
Main occupation unpaid domestic work	2	0.90	0.89
N	2463		99

Table E1: 20-year differential treatment effects for the early childhood cohort by birth-year

		Born in 1997		Born in 1998		Born in 1999		P-value
		Obs	Coef.	Obs	Coef.	Obs	Coef.	1997=1998=1999
		(1)		(2)		(3)		(4)
Completed primary school								
	T1	1644	0.017	1365	0.02	1077	0.017	0.938
	s.e.		(0.013)		(0.013)		(0.018)	
	mean T2		0.93		0.932		0.939	
Completed lower secondary school								
	T1	1644	0.038	1365	0.061	1077	0.069	0.497
	s.e.		(0.026)		(0.03)**		(0.036)*	
	mean T2		0.749		0.708		0.711	
Completed upper secondary school								
	T1	1644	0.038	1365	0.012	1077	0.076	0.148
	s.e.		(0.032)		(0.029)		(0.022)***	
	mean T2		0.321		0.274		0.118	
Completed some tertiary education								
	T1	1644	0.021	1365	0.023	1077	0.004	0.267
	s.e.		(0.012)*		(0.009)***		(0.007)	
	mean T2		0.046		0.016		0.009	
Grades attained (years of education)								
	T1	1644	0.327	1365	0.325	1077	0.429	0.855
	s.e.		(0.183)*		(0.199)		(0.213)**	
	mean T2		9.291		9.117		8.93	
Last schooling in locality of origin								
	T1	1493	-0.084	1241	-0.075	1010	-0.05	0.756
	s.e.		(0.035)**		(0.037)**		(0.044)	
	mean T2		0.394		0.393		0.388	
Still studying								
	T1	1664	0.032	1372	0.053	1084	-0.015	0.25
	s.e.		(0.02)		(0.024)**		(0.04)	
	mean T2		0.127		0.178		0.383	
Log(expected income at age 30)								
	T1	1289	0.065	1111	0.156	908	0.005	0.247
	s.e.		(0.055)		(0.066)**		(0.065)	
	mean T2		11.64		11.601		11.732	
Expected income at age 30 - trimmed at 2%								
	T1	1269	9080.091	1090	18112.1	886	8823.263	0.709
	s.e.		(7681.677)		(9462.654)*		(8947.917)	
	mean T2		141582.7		140914		151681.3	

Notes: ITT estimates with region and village population size fixed effects. Col. 1 shows results for children born in 1997, Col. 2 those born in 1998 and Col. 3t those born in 1999. Col. 4 shows the p-values for tests that the effects are identical between children with different birth years. Each column shows the number of observations and the estimated differential treatment effect in the first row, standard errors are in the second cell row parentheses, clustered at locality level, with ***, **, * indicating significance at 1, 5 and 10%; and mean in T2 localities in the third row.

Table E2: 20-year differential treatment effects for the early childhood cohort: income

		All		Women		Men		P-value
		Obs	coef.	Obs	coef.	Obs	coef.	Women=Men
		(1)		(2)		(3)		(4)
Labor income last 12 months								
	ITT	3714	5077.113	1947	600.098	1767	7823.468	0.094*
	s.e.		(2255.386)**		(1972.408)		(4070.606)*	
	Mean control		36781.11		20209.85		56281.14	
Has US income								
	ITT	3758	0.003	1956	0.002	1802	0.003	0.937
	s.e.		(0.004)		(0.002)		(0.009)	
	Mean control		0.012		0.001		0.025	
Has labor income								
	ITT	3758	0.028	1956	0.004	1802	0.031	0.507
	s.e.		(0.019)		(0.026)		(0.026)	
	Mean control		0.546		0.352		0.769	
Log(labor income last 12 months), conditional on working								
	ITT	2100	0.034	695	-0.031	1405	0.058	0.216
	s.e.		(0.041)		(0.062)		(0.05)	
	Mean control		10.904		10.78		10.971	

Notes: ITT estimates with region and village population size fixed effects. Income includes labor income in Mexican pesos of the 2 main occupations in the last 12 months, trimmed at 1%, and with dollar income of US residents converted to Mexican pesos using PPP-adjusted exchange rate. Col. 1 shows results for the early childhood cohort, Col. 2 for female and Col. 3 for males. Col. 4 shows the p-values for tests that the effects are identical between males and females. Each column shows the number of observations and the estimated differential treatment effect in the first row, standard errors are in the second cell row parentheses, clustered at locality level, with ***, **, * indicating significance at 1, 5 and 10%; and mean in T2 localities in the third row.

Table E3: 20-year differential treatment effects for the early childhood cohort: early fertility and health behaviors

		All		Women		Men		P-value
		Obs	coef.	Obs	coef.	Obs	coef.	Women=Men
		(1)		(2)		(3)		(4)
Ever married or in civil union								
	ITT	4137	-0.009	2083	-0.022	2054	0.014	0.163
	s.e.		(0.016)		(0.022)		(0.018)	
	Mean control		0.28		0.385		0.167	
Has any children								
	ITT	4115	-0.007	2074	-0.011	2041	0.005	0.499
	s.e.		(0.015)		(0.025)		(0.014)	
	Mean control		0.216		0.321		0.103	
First child by age 18								
	ITT	4106	-0.009	2070	-0.01	2036	0	0.669
	s.e.		(0.014)		(0.023)		(0.011)	
	Mean control		0.162		0.257		0.06	
Is household head or spouse of hh head								
	ITT	4101	0.011	2070	0.012	2031	0.014	0.911
	s.e.		(0.014)		(0.02)		(0.018)	
	Mean control		0.186		0.231		0.138	
Smokes (last 12 months)								
	ITT	3831	0.01	1968	0.014	1863	-0.01	0.467
	s.e.		(0.015)		(0.01)		(0.027)	
	Mean control		0.17		0.029		0.328	
Drinks alcohol (last 12 months)								
	ITT	3831	0.025	1968	0.014	1863	0.017	0.795
	s.e.		(0.017)		(0.016)		(0.025)	
	Mean control		0.264		0.099		0.451	

Notes: ITT estimates with region and village population size fixed effects. Col. 1 shows results for the early childhood cohort, Col. 2 for female and Col. 3 for males. Col. 4 shows the p-values for tests that the effects are identical between males and females. Each column shows the number of observations and the estimated differential treatment effect in the first row, standard errors are in the second cell row parentheses, clustered at locality level, with ***, **, * indicating significance at 1, 5 and 10%; and mean in T2 localities in the third row.

Table E4: 20-year differential treatment effects for the school cohort: other occupations

	All		Women		Men		P-value
	Obs	coef.	Obs	coef.	Obs	coef.	Women=Men
	(1)		(2)		(3)		(4)
Main occupation in non-agricultural sector							
T1	2152	-0.015	1096	-0.008	1056	-0.028	0.57
s.e.		(0.025)		(0.03)		(0.036)	
mean T2		0.48		0.323		0.645	
Main occupation in non-agricultural non-skilled job							
T1	2152	-0.023	1096	-0.007	1056	-0.044	0.37
s.e.		(0.024)		(0.027)		(0.035)	
mean T2		0.393		0.25		0.543	
Main occupation in agricultural self-employment							
T1	1882	0.007	1002	-0.01	880	0.02	0.25
s.e.		(0.013)		(0.006)		(0.025)	
mean T2		0.06		0.013		0.118	
Main occupation in agricultural wage employment							
T1	1882	0.014	1002	0.002	880	0.012	0.928
s.e.		(0.016)		(0.009)		(0.031)	
mean T2		0.107		0.018		0.214	
Main occupation in non-agricultural self-employment							
T1	1882	0.012	1002	-0.003	880	0.025	0.449
s.e.		(0.014)		(0.016)		(0.025)	
mean T2		0.081		0.059		0.109	
Main occupation in non-agricultural wage-employment							
T1	1882	-0.022	1002	0.004	880	-0.068	0.113
s.e.		(0.024)		(0.027)		(0.037)*	
mean T2		0.381		0.253		0.537	
Main occupation in agricultural or commerce self-employment							
T1	1883	0.015	1002	-0.01	881	0.038	0.086*
s.e.		(0.014)		(0.01)		(0.025)	
mean T2		0.076		0.028		0.134	
Occupation in agricultural or commerce self-employment - first and second job							
T1	1884	0.017	1003	-0.011	881	0.044	0.049**
s.e.		(0.014)		(0.01)		(0.026)*	
mean T2		0.081		0.031		0.143	

Notes: ITT estimates with region and village population size fixed effects. Col. 1 shows results for the schooling cohort, Col. 2 for female and Col. 3 for males. Col. 4 shows the p-values for tests that the effects are identical between males and females. Each column shows the number of observations and the estimated differential treatment effect in the first row, standard errors are in the second cell row parentheses, clustered at locality level, with ***, **, * indicating significance at 1, 5 and 10%; and mean in T2 localities in the third row.

Table E5: 20-year differential treatment effects for the school cohort: job characteristics

		All		Women		Men		P-value
		Obs	coef.	Obs	coef.	Obs	coef.	Women=Men
		(1)		(2)		(3)		(4)
Receives social security benefits from job								
	ITT	1882	-0.011	1002	-0.003	880	-0.034	0.419
	s.e.		(0.02)		(0.019)		(0.034)	
	Mean control		0.184		0.097		0.289	
Receives pay every 15 days								
	ITT	1884	-0.003	1003	0.004	881	-0.015	0.534
	s.e.		(0.015)		(0.014)		(0.025)	
	Mean control		0.098		0.061		0.143	
Receives weekly pay								
	ITT	1884	-0.016	1003	-0.014	881	-0.048	0.442
	s.e.		(0.024)		(0.026)		(0.035)	
	Mean control		0.429		0.235		0.665	
Number of months worked - 1&2 job combined								
	ITT	1884	0.086	1003	-0.285	881	0.07	0.516
	s.e.		(0.276)		(0.358)		(0.201)	
	Mean control		6.888		3.847		10.581	
Worked 12 months in same job								
	ITT	1884	-0.001	1003	-0.053	881	0.03	0.074*
	s.e.		(0.025)		(0.029)*		(0.033)	
	Mean control		0.48		0.292		0.708	

Notes: ITT estimates with region and village population size fixed effects. Col. 1 shows results for the schooling cohort, Col. 2 for female and Col. 3 for males. Col. 4 shows the p-values for tests that the effects are identical between males and females. Each column shows the number of observations and the estimated differential treatment effect in the first row, standard errors are in the second cell row parentheses, clustered at locality level, with ***, **, * indicating significance at 1, 5 and 10%; and mean in T2 localities in the third row.

Table E6: 20-year differential treatment effects for the school cohort: partner characteristics and health behaviors

		All		Women		Men		P-value
		Obs	coef.	Obs	coef.	Obs	coef.	Women=Men
		(1)		(2)		(3)		(4)
Selection/information available for education partner								
	ITT	2153	-0.005	1097	0.004	1056	-0.013	0.646
	s.e.		(0.022)		(0.029)		(0.029)	
	Mean control		0.716		0.729		0.702	
Education(grades attained) of partner								
	ITT	1531	0.161	799	-0.003	732	0.343	0.235
	s.e.		(0.184)		(0.236)		(0.247)	
	Mean control		8.37		8.21		8.544	
Education(grades attained) higher than partner								
	ITT	1529	0.018	798	0.059	731	-0.027	0.059*
	s.e.		(0.024)		(0.035)*		(0.033)	
	Mean control		0.329		0.347		0.309	
Education(grades attained) lower than partner								
	ITT	1529	-0.01	798	-0.023	731	0.007	0.437
	s.e.		(0.025)		(0.033)		(0.035)	
	Mean control		0.297		0.276		0.319	
Age partner								
	ITT	1580	0.033	822	-0.125	758	0.252	0.498
	s.e.		(0.314)		(0.416)		(0.37)	
	Mean control		31.87		34.738		28.727	
Age 2 or more years higher than partner								
	ITT	1582	0.028	822	0.018	760	0.04	0.52
	s.e.		(0.025)		(0.023)		(0.038)	
	Mean control		0.333		0.106		0.58	
Age 2 or more years lower than partner								
	ITT	1582	-0.003	822	-0.018	760	0.009	0.698
	s.e.		(0.024)		(0.02)		(0.039)	
	Mean control		0.719		0.922		0.498	
Partner from a different locality - cond								
	ITT	1482	0.017	786	0.044	696	-0.019	0.188
	s.e.		(0.029)		(0.036)		(0.039)	
	Mean control		0.661		0.65		0.674	
Partner participates in labor market								
	ITT	1629	-0.006	840	-0.001	789	-0.015	0.788
	s.e.		(0.026)		(0.011)		(0.027)	
	Mean control		0.591		0.981		0.179	
Smokes (last 12 months)								
	ITT	1884	0.02	1004	-0.004	880	0.035	0.204
	s.e.		(0.017)		(0.011)		(0.032)	
	Mean control		0.142		0.036		0.27	
Drinks alcohol (last 12 months)								
	ITT	1884	0.016	1004	0.036	880	-0.031	0.152
	s.e.		(0.024)		(0.019)*		(0.035)	
	Mean control		0.303		0.066		0.59	

Notes: ITT estimates with region and village population size fixed effects. Col. 1 shows results for the schooling cohort, Col. 2 for female and Col. 3 for males. Col. 4 shows the p-values for tests that the effects are identical between males and females. Each column shows the number of observations and the estimated differential treatment effect in the first row, standard errors are in the second cell row parentheses, clustered at locality level, with ***, **, * indicating significance at 1, 5 and 10%; and mean in T2 localities in the third row.

Table E7 : 20-year differential treatment effects for school cohort: heterogeneity based on lower secondary school in village

	Completed lower secondary school	Has US income	Labor income last 12 months	Log (labor income last 12 months, cond. on working)	Living in locality of origin
	(1)	(2)	(3)	(4)	(5)
T1	0.109***	0.0482***	13279**	0.219***	-0.0778**
	(0.036)	(0.014)	(5735)	(0.070)	(0.032)
T1*Secondary school in village	-0.124**	-0.0216	-14866	-0.195	0.0924
	(0.053)	(0.029)	(11633)	(0.12)	(0.059)
Secondary school in village	0.226***	0.0472**	16394*	0.290***	-0.0905*
	(0.043)	(0.021)	(9395)	(0.11)	(0.052)
Mean T2	0.606	0.044	57045	11.02	0.386
Obs	2141	1834	1801	1187	2153
P-value T1+T1*secondary in village	0.684	0.290	0.875	0.815	0.769

Notes: ITT estimates with region and village population size fixed effects. "Secondary school in village" is a binary variable indicating whether there was a lower secondary school located in the village at baseline. Standard errors in parentheses, clustered at locality level, with ***, **, * indicating significance at 1, 5 and 10%.