Experimental Evidence from a Conditional Cash Transfer Program: Schooling, Learning, Fertility, and Labor Market Outcomes After 10 Years

By TANIA BARHAM, KAREN MACOURS AND JOHN A. MALUCCIO[1]

August 2023

Abstract: Conditional cash transfer programs are the anti-poverty program of choice in many developing countries, aiming to improve human capital and break the intergenerational transmission of poverty. A decade after a randomized 3-year CCT program began, earlier exposure during primary school ages when children were at risk of dropout led to higher labor market participation for young men and women and higher earnings for men. Results highlight the roles of the different program components with variation in timing of access to nutrition, health and education investments translating into substantial differential effects on learning for men and reproductive health outcomes for women.

JEL Codes: I25, I38, I18, I28, J13
Key words: CCT, long-term effects, labor markets, education, learning, fertility, age of menarche, Nicaragua

[1] Barham: Department of Economics and IBS, University of Colorado at Boulder, Boulder, CO 80309-0256 (tania.barham@colorado.edu); Macours: Paris School of Economics and INRA, 48 Boulevard Jourdan, 75014 Paris, France (karen.macours@psemail.eu); Maluccio, Department of Economics, Middlebury College, Middlebury VT 05753 (maluccio@middlebury.edu). This paper combines and extends CEPR Discussion Papers CEPR 11937 and 13165. Acknowledgments: This research would not have been possible without the support of Ferdinando Regalia of the Inter-American Development Bank (IDB). We gratefully acknowledge generous financial support from IDB, the Initiative for International Impact Evaluation (3ie: OW2.216) and the National Science Foundation (SES 11239945 and 1123993). See also AEA RCT registry AEARCTR-0001572. We are indebted to Veronica Aguilera, Enoe Moncada and the survey team from CIERUNIC for excellent data collection and for their persistence in tracking. We also acknowledge members of the *Red de Protección Social* program team (in particular Leslie Castro, Carold Herrera and Mireille Vijil) for discussions regarding this research and Emma Sanchez Monin for facilitating the data collection process. We thank Teresa Molina-Millán, Olga Larios, Jana Parsons and Gisella Kagy for help with data preparation and Vincenzo di Maro for numerous contributions to the 2010 survey. Finally, we are grateful for comments received from Guillermo Cruces, Pierre Dubois, Esther Duflo, Berk Özler, Norbert Schady, several anonymous referees and others during presentations at the IDB, Northeast Universities Development Conference 2012, Allied Social Sciences Association 2013, Colby College, Middlebury College, European Development Network Conference 2013, Population Association of America meetings, UCL, Louvain-la-Neuve, Pompeu Fabra and at the Impact Evaluation Network conference of LACEA. All remaining errors and omissions are our own.

I. Introduction

Interventions aimed at increasing the nutrition, health and education of children are often motivated by the possibility of breaking the intergenerational transmission of poverty. Theory suggests that investments in child human capital can improve future economic outcomes, for example through higher incomes in adulthood. There is substantial evidence that a variety of interventions can increase human capital in low- and middle-income countries in the short term, but less is known about whether they live up to their promise in the longer term.[2] There is also little evidence on the potential importance of specific timing of interventions in late childhood.

A prominent example in this class of interventions is the conditional cash transfer (CCT) program. Started in 1997 in Mexico and Brazil, CCTs spread to 60+ countries worldwide and covered 25% of Latin America by 2013 (World Bank 2015). Numerous rigorous evaluations demonstrate short-term effects including improvements in nutrition and health for young children, increases in schooling for older children and reductions in poverty for households (Fiszbein and Schady 2009; Bastagli et al. 2016). Evidence on longer-term effects is more mixed (Molina-Millán et al. 2019). Using the randomized phase-in of a CCT, we provide new evidence on how experimental differences in the specific timing of CCT exposure during primary school ages can have long-term differential effects on education, health and labor market outcomes.

More specifically, we exploit the randomized phase-in and eligibility rules of a CCT in rural Nicaragua to estimate differential intent-to-treat (ITT) effects 10 years after the start of the program. All households in the early treatment group were eligible for transfers for three years from late 2000 to late 2003; those in the late treatment group were also eligible for transfers for three years, but from 2003 to 2005. In both treatments, all households were eligible for conditional transfers for nutrition and health, and households with children 7–13 years old who had not completed fourth grade were additionally eligible for conditional transfers for education. The program also increased exposure to reproductive health information and services. Together, the conditional transfers provided incentives and resources for the children to remain in school, as well as means to improve food availability and nutrition for them.[3]

[2] Important exceptions include evidence on long-term effects of family planning and early childhood health interventions (Barham et al. 2018), early childhood stimulation (Gertler et al. 2014), early childhood nutrition (Hoddinott et al. 2008), deworming interventions (Baird et al. 2016), education subsidies and HIV prevention education (Duflo et al. 2015), and school vouchers (Bettinger et al. 2016).

[3] Short-term evaluation results confirm that the CCT led to relatively large increases in enrollment, grade completion and household food expenditures (Maluccio and Flores 2005). The impacts parallel those of many other first

To examine the long-term effects, we collected data in 2010 on individuals and households first interviewed at baseline in 2000, prior to randomization. Consequently, the follow-up survey measured outcomes approximately 10 years after the program start, which was seven years after households in the early treatment group had stopped receiving transfers. At that point individuals 7–13 years old at baseline were transitioning into early adulthood (17–23 years old) with the majority in their initial years in the labor market. Evidence during this labor market entry period is valuable for understanding the mechanisms underlying the long-term effects of CCTs, particularly if those initial experiences can set individuals on different labor market trajectories.

Given the potential relationship between CCT program exposure and migration for the young adults, we tracked all of them throughout Nicaragua and to Costa Rica, the main international migration destination. International tracking is rare in these types of studies. Final attrition is between 4–22% (depending on the outcome and sample), is balanced between early and late treatment, and yields analysis samples balanced on baseline observables. We consider various estimation strategies to gauge the extent of possible remaining selectivity.

To help understand how the CCT-related investments influenced education, health and labor market outcomes we posited a dynamic human capital production function model (Cunha and Heckman 2007) with three periods corresponding to the two program implementation phases described above and to a third period ending with the follow-up survey in 2010. The possibility of self-productivity, dynamic complementarities and, in particular, sensitive periods for education and health investments during primary school ages helps us interpret how the experimental variation in timing led to longer-term differentials in outcomes.

We find that earlier exposure to the CCT starting at primary school ages when children were at risk of dropout (9–12 years old) increased labor market participation and earnings a decade later when the beneficiaries were young adults. To understand what underlies the gains, we examine several intermediate human capital outcomes along the causal pathway. Much of the CCT literature examining effects on primary school age children focuses on education. But there are other mechanisms through which CCTs can affect beneficiaries and all program components—rather than just the education component—potentially influence later outcomes. In

generation CCT programs that, like the Nicaraguan program we study, promoted nutrition, health and education, and targeted children over a relatively wide age range. Recent CCT programs, in contrast, often focus more narrowly on one specific objective (e.g., education or health) or age group.

addition to education, this paper spotlights the role of improved food availability and nutrition during the preteen years which can influence health and, for girls, the timing of menarche and subsequent sexual and reproductive health outcomes. This is particularly relevant for understanding the effects of differential timing of exposure for girls since their education, fertility and labor market outcomes are closely linked. We therefore examine results by sex.

For both men and women separately, earlier exposure at primary school ages (9–12 years old) increased labor market participation. Differential ITT effects on intermediate outcomes by sex, however, reveal salient differences in mechanisms resulting from the timing of the human capital investments and possibly different sensitive periods, interaction with labor market opportunities and the role of fertility for labor market outcomes of young adults in Nicaragua.

By the 10-year follow-up most young men had completed their schooling and virtually all were working. The differential ITT estimates indicate that earlier exposure to the CCT at ages when they were likely to drop out of school led to sustained schooling and learning gains of nearly 0.2 standard deviations (SD). In addition, they were more likely to have wage work and migrate temporarily for higher paid jobs, earning ~15% more per month worked. The pattern of results for young men is in line with standard human capital theory in which investment during sensitive periods leads to more schooling, more learning and subsequent changes in labor market outcomes including higher earnings. In this regard, results for males provide relatively straightforward evidence on the long-term returns to CCTs and on the potential for CCTs to reduce the intergenerational transmission of poverty through improved education.

For the young women we also estimate positive and significant differential ITT effects on schooling; however, despite short-term evidence of gains in early treatment there are no differential effects on learning. This can be explained in part by prior enrollment patterns. Girls in rural Nicaragua typically remain in school longer than boys during the early teenage years, and this may have enabled late treatment girls to catch up later when their households were benefitting from the CCT, even though they themselves were not directly exposed to the education component. Results also indicate that relative to late treatment, average age of menarche increased for early treatment girls and they became sexually active later. These differences can be linked to the random variation in the timing of CCT transfer-related nutrition shocks occurring during ages important for physical development. Consistent with their older age of menarche, early treatment girls started childbearing later and had lower body mass compared to those in the late treatment. Taken together, the pattern of results for females suggests the long-term effects on their labor market outcomes reflect CCT induced changes in nutrition and related reproductive health outcomes, in addition to changes in schooling.

The findings directly relate to the literature on long-term impacts of exposure to CCTs during school-going ages, including Behrman et al. (2009a, 2011), Barrera-Osorio et al. (2019), Baird et al. (2019), Cahyadi et al. (2020) and Molina-Millán et al. (2020).[4] Prior research largely finds that CCTs lead to higher schooling, but is less conclusive for other outcomes including learning and, with the exception of Parker and Vogl (2018) and Araujo and Macours (2021) for Mexico, there is little evidence on labor market or income gains. There is also no evidence on whether and how the specific timing of the interventions changes program impacts.

Our paper contributes to this literature in several ways. We organized extensive tracking of migrants (resulting in relatively low attrition rates) and collected rich information on the diverse portfolio of economic activities in which young Nicaraguan adults engage. This enabled a comprehensive examination of longer-term impacts on the labor market, which identified temporary migration for work as a path to increasing earnings. Measurement of a wide set of intermediate outcomes allowed us to explore effects along the causal pathway to young adult labor market and earning outcomes. That analysis provides a detailed picture of the ways in which CCT exposure during primary-school age years can lead to longer-term outcomes including the importance of timing of investments and how patterns can differ by sex. Finally, we also found that differential exposure to a CCT affected child learning in this low-income context, in line with some results in the literature (Baird et al. 2011; Duque et al. 2018), but in contrast to other results (Filmer and Schady 2014; Baird et al. 2019).

The findings also relate to literature examining the links between schooling, fertility, marriage and the labor market for women. Research on age of menarche and schooling in India and Bangladesh demonstrates that later menarche is associated with higher schooling and delayed marriage, due at least in part to cultural norms (Field and Ambrus 2008; Khanna 2020). Research on the longer-term impacts of scholarships and educational subsidies reveals that effects on fertility-related outcomes are often directly linked to the educational incentives;

[4] The findings also relate to work examining longer-term impacts on individuals of CCT exposure earlier in childhood (Behrman et al. 2009b; Fernald et al. 2009; Barham et al. 2013; Araujo et al. 2018), as well as impacts on household economic well-being of exposure to CCT or other cash transfer programs (Gertler et al. 2012; Macours and Vakis 2016; Banerjee et al. 2016; Bandiera et al. 2017; Handa et al. 2018; Haushofer and Shapiro 2018).

inducing girls to remain in school longer can postpone childbearing (Baird et al. 2011; Filmer and Schady 2014; Duflo et al. 2015, 2021). Financial incentives for delaying marriage given to teenage girls (initially) not in school in Bangladesh actually increased schooling while reducing marriage and teen childbearing, although empowerment training for them only resulted in higher self-employment (Buchmann et al. 2018). Reproductive health and vocational trainings for adolescent girls in Uganda similarly led to increases in self-employment but also to reductions in teen pregnancy and marriage (Bandiera et al. 2020).

Our paper contributes to this literature by emphasizing the nutrition and health related component of transfer programs as an additional channel through which female fertility outcomes—important for labor market outcomes—can be influenced.[5] We document that timing of exposure to the nutritional and reproductive health program components can affect the age of menarche and subsequent fertility outcomes. One reason this is an important evidence gap to fill is that the medical and nutritional literatures indicate exposure to transfers (and their likely effects on household spending and food availability) at different ages can have implications for the onset of puberty and therefore affect later fertility and other outcomes. Furthermore, research indicates that biological sexual maturation may be particularly sensitive to nutritional shocks for young adolescent girls who experienced poor nutritional status earlier in life (INSERM 2007), as was almost certainly the case for our study population.[6] Consequently, a fuller understanding of the potential differential effects of timing of exposure to the nutrition components, which are common to most large CCT programs but have received relatively little attention in comparison to exposure to the education components, is critical for optimal policy design.

II. The Nicaraguan CCT Program and Experimental Design

The *Red de Protección Social* was a CCT implemented by the Nicaraguan government to address both current and future poverty among rural households. Transfers were made every two months to female caregivers and averaged 18% of total pre-program household expenditures. The transfers came with formal enforced conditionalities alongside a strong social marketing message that the money was meant for food, health and education expenditures. A household

[5] As such, the evidence also complements studies examining the shorter-term reproductive health and fertility effects of related interventions, most of which focus on adult program beneficiaries (Stecklov et al. 2007; Lamadrid-Figueroa et al. 2010; Avitabile 2012; Todd et al. 2012). See also Khan et al. (2016) for a systematic review.

[6] See Appendix H for a review of the relevant medical and nutrition literature.

representative signed an agreement they would comply with the conditions and spend the money as intended, although actual expenditures were not monitored or part of the formal conditions.

Like most first-generation CCTs, there were two core components: health and education.

The first component aimed to improve food security, nutrition and health via a universal cash transfer to all households, regardless of size or composition. This nutrition and health transfer was paid bimonthly and conditional on preventive healthcare visits for any children under five years old and on attendance by the caregiver at monthly health education workshops. Local supply of the required health services was increased in parallel with the program.

The second component aimed to improve education. All households with a child 7–13 years old who had not completed fourth grade were eligible. The household received a fixed bimonthly school attendance cash transfer, about half the size of the nutrition and health transfer and conditional on enrollment and regular attendance of all eligible children. For each eligible child, the household also received a small annual cash transfer intended for school supplies. We refer to the combined school attendance and school supplies transfers as the education transfer.

Although modeled largely after Mexico's *PROGRESA*, there were two important differences in the Nicaraguan CCT. First, in Nicaragua beneficiary households were only eligible to receive the program for a fixed three-year period, after which it was not possible to renew. Second, the Nicaragua CCT was focused on early dropout and limited eligibility for education transfers to the first four grades of primary. Household decisions on timing of child schooling and investments in human capital (and any consequent effects), therefore, might differ compared to decisions made under a program like *PROGRESA* with continuing eligibility and without grade limits.

A randomized evaluation was incorporated into the design of the CCT in six rural municipalities from three regions in central and northern Nicaragua, chosen based on their extreme poverty and relatively low health and educational outcomes. In these municipalities, 42 of 59 rural localities were selected based on a locality-level marginality index.[7] A program registry census of all households in the 42 localities was done in May 2000 prior to randomization. Localities were ordered by the marginality index and divided into seven strata of six localities each, and randomization by strata done publicly in July 2000 with 21 localities randomized to the early treatment group and 21 to late treatment. Households in early treatment

[7] Localities were defined as census *comarcas* from the 1995 Nicaraguan national census. These were administrative areas including as many as 10 small villages and averaging approximately 250 households per locality.

localities received their first transfers in November 2000 and were eligible for up to three years of transfers, with the last delivered in late 2003. After that they received no further transfers, nor were they subject to any conditions. After randomization, households in late treatment were informed that the program would start later in their localities; they were incorporated in January 2003 and likewise eligible to receive up to three years of transfers. Prior notification introduces the possibility of anticipation effects (Section III). In December 2005, the program ended.[8]

Previous short-term analysis demonstrates the randomized sample was balanced at baseline (Maluccio and Flores 2005). Appendix Table A1 demonstrates balance for the samples in this paper. Compliance with the experimental design was high; ~85% of households took up at least one component of the program and there was negligible contamination of the late treatment.

Apart from timing, there were some differences between the early and late treatment program potentially important for our study. First, scheduled transfer amounts were modestly lower for late treatment households, approximately 88% of the early treatment amount. New services were added, however, arguably offsetting this reduction in transfer amounts. Starting in 2003, when the late treatment group became eligible for the program, modern contraception and prenatal consultations were made available to all beneficiaries through the healthcare providers. In addition, as part of the conditions for the nutrition and health transfer, adolescents were now required to attend workshops covering among other things sexual and reproductive health.[9] The workshops were available to both groups, but after 2003 attendance was a condition only for the late treatment. Correspondingly, these services were less intensively implemented and taken up in early treatment localities. Apart from these differences, the eligibility criteria and principal services and conditions were the same across the early and late treatment groups (Appendix C).

III. Conceptual Framework and Identification

A. Conceptual Framework

To guide interpretation of the long-term effects of the differential timing of the CCT we outline a multi-period model of human capital formation beginning in late childhood building on

[8] Nationally, all Nicaraguan government CCT programs ended in 2006 having benefitted ~30,000 households from 2000–6. As such, general equilibrium effects on labor or other markets are likely to have been limited.

[9] Comparable topics were covered with adult caregivers in the health workshops required throughout the program in both treatment groups. Most first generation CCTs included similar workshops (Fiszbein and Schady 2009).

Cunha and Heckman (2007) and Attanasio et al. (2020).[10] Reflecting the main program components, we posit two dimensions of human capital important for labor market outcomes: education and health. The model has three periods. The first two correspond to program exposures in the early (2000–3) and late (2003–5) experimental treatment groups, and the third corresponds to the years after the program had ended until our final measurement in the 2010 long-term evaluation survey (2006–10). The evolution of education and health human capital are characterized using dynamic production functions:

$$\theta_{t+1}^{e} = f_t^{e}(\theta_t^{e}, \theta_t^{h}, I_t, Z_1, a_1) \quad \text{(1a)}$$

$$\theta_{t+1}^{h} = f_t^{h}(\theta_t^{e}, \theta_t^{h}, I_t, Z_1, a_1) \quad \text{(1b) with } t = 1, 2, 3$$

Individuals are first observed in period 1 at age a_1 with previously accumulated education (θ_1^e) and health (θ_1^h) human capital levels and background characteristics (Z_1) such as sex. Public and private investments in each period *t* are represented by $I_t = (I_t^{pub}, I_t^{priv})$.

Under conditions outlined in Cunha and Heckman (2007), the dynamic framework points to the potential for self-productivity, dynamic complementarity and sensitive periods in human capital production. Self-productivity refers to the possibility that for a given level of investment higher levels of human capital in one period give rise to higher human capital in the next, both within and across types, i.e., $\partial\theta_{t+1}^{j}/\partial\theta_t^{k} \equiv \partial f_t^{j}(\theta_t^{e}, \theta_t^{h}, I_t, Z_1, a_1)/\partial\theta_t^{k} > 0$ for $j, k \in \{e, h\}$. Dynamic complementarity refers to the possibility that the marginal productivity of investment increases with human capital levels in period *t*, i.e., $\partial^2\theta_{t+1}^{j} / \partial\theta_t^{k}\partial I_t \equiv \partial^2 f_t^{j}(\theta_t^{e}, \theta_t^{h}, I_t, Z_1, a_1)/\partial\theta_t^{k}\partial I_t > 0$. Last, sensitive periods refer to the possibility that a given level of investment has a higher return in some periods. t^* is a sensitive period if $\partial\theta_{t+1}^{j}/\partial I_{t*} > \partial\theta_{t+1}^{j}/\partial I_s$ holds for periods $s \neq t^*$. The relevance and strength of all three phenomena likely differ by human capital type, period and sex.

The CCT operates in early treatment localities in period 1, late treatment localities in period 2 and no longer operates in period 3. Therefore, it experimentally alters the timing of program-related public investment (I_t^{pub}) from period 1 to period 2. Changes in public investment directly affect human capital but can also lead to changes in private investment (I_t^{priv}) as households comply with both hard and soft program conditions and optimize over time. For example, early

[10] Although there are some important exceptions (including Cunha et al. 2010 and Carneiro et al. 2021), the late childhood period we begin with is less well studied in the dynamic human capital production function literature.

treatment households who experience increased public investment in period 1 can increase their private investment in period 2 to benefit from potential dynamic complementarities. Conversely, advance notification of future (period 2) public investment for late treatment households can lead to anticipation effects in the form of greater period 1 private investment by them to benefit from the potential dynamic complementarities when the public investment arrives. The extent of the private investment response also depends on the degree of self-productivity and relative sensitivity of investment in periods 1 and 2, as well as the resources available to households.

These program-influenced investment paths, different for early versus late treatment, go on to affect the evolution and final levels of human capital. Notably, the existence of sensitive periods means that even with positive self-productivity and dynamic complementarities, earlier investment is not necessarily more productive in the longer term, for example if later investment is made in more sensitive periods. Hence, the conceptual model alone cannot be used to determine whether period 1 or 2 public investment is more beneficial for human capital production, making long-term differentials resulting from early versus late treatment theoretically ambiguous. In Section III.B we draw on program design, pre-program human capital (in particular school enrollment patterns and early childhood nutritional status) and the medical literature on sexual maturation for girls to explore which periods might be more sensitive to investments in education and health. The randomized timing of the CCT together with the fixed 3-year exposure window offer an opportunity to examine how different investment shocks during distinct late childhood periods with potentially different sensitivities can affect human capital. Other CCT programs with the possibility of renewed or ongoing eligibility would likely lead to different results given dynamic complementarities and self-productivity.

Last, regardless of which treatment has larger impacts on human capital, improvements in either education or health in periods 1 and 2 can affect labor market outcomes in period 3. Improvements in education and health could both improve labor market outcomes but it is also possible that they could have offsetting effects on labor market outcomes, for example if improved health leads to earlier fertility that in turn lowers labor force participation.

B. Identification of 2010 Long-term Experimental Effects of the Differential Timing of Exposure and Cohorts of Interest

The randomized timing of program implementation generated variation in the ages at which children were exposed to the CCT. Because early and late experimental treatment groups both received the CCT and there is no pure control group, we can estimate differentials rather than absolute program effects. For outcomes in which early and late treatments both have positive absolute effects, the differential will be smaller than the absolute effect in the early treatment. This includes the possibility of similarly sized absolute effects in both treatment groups that offset one another and result in no net differential.

Following the original experimental design, we first examine children 7–12 years old at the start of the program in November 2000, i.e., at baseline. The 7–12 cohort is the largest cohort directly exposed to the education component from the start in early treatment for whom there is detailed survey information in 2010.[11] We then limit to 9–12-year-olds at baseline to focus on periods we expect to be more sensitive to investment in education or health, and consequently for which larger long-term differentials in labor market outcomes might be observed. Limiting the age range to the 9–12 cohort leads to larger potential differences between treatment groups in direct exposure to the education component. This is both because it excludes the younger 7–8 cohort who were potentially exposed to the full three years of the program in either early or late treatment and because older children in the 9–12 cohort would have reached 13 and therefore aged out of eligibility for the education component in late treatment before the CCT became available to them.[12] The more limited 9–12 range also reduces overlap in ages exposed across the early and late treatment groups, sharpening the contrast between the experimental groups.

To understand which period was potentially more sensitive for investment in education in this context, we examine schooling patterns. Investment at ages when children are already enrolled in school and likely to continue, or after they have already been out of school for some time, will likely be less effective than investment at ages when children are still enrolled but at high risk of dropping out (de Janvry and Sadoulet 2006). Baseline enrollment rates in program

---

[11] We do not include 13-year-olds in the analysis because they were not tracked beyond their original households nor administered the individual-level survey due to budgetary constraints and their limited (at most one year) potential direct exposure to the CCT education component.

[12] Appendix Table C1 summarizes differential exposure by age for study cohorts. Child age mattered for eligibility for the education component but did not matter for exposure to the nutrition and health component for which all households were eligible. Therefore, the CCT could have affected the education of children outside the explicitly targeted ages or grades for the education component if they benefited directly from the nutrition and health transfer. They also could have benefited indirectly from education transfers made to other eligible children in the household.

areas indicate that the risk of school dropout increased sharply for boys starting at about age 11 and for girls at age 13. Boys in the 9–12 cohort in the late treatment group had turned 11–14 years old by 2003 when the CCT began for them and were thus more likely to have already dropped out, with the consequence that for them the CCT might have started too late to affect their schooling. By similar reasoning, the CCT also might have started too late to substantially benefit late-treatment girls 11–12 years old at baseline, as they would have reached 13–14 by 2003. Combined with the age-related program eligibility rules, the dropout patterns make it likely that the early treatment period (or period 1) was more sensitive to education investment, although to a lesser extent for girls because of their older typical dropout ages. Given that the early treatment group benefited from the public investment in period 1, all else equal we expect positive long-term differentials on education human capital, especially for boys.

To understand which period was potentially more sensitive for investment in health, we consider the implications of the CCT providing resources intended for food via the nutrition and health component; children in the 9–12 cohort in beneficiary households likely experienced improved nutrition during the program. Although important for male development, nutrition at these ages is particularly relevant for girls because of its potential effect on sexual maturation.

The nutrition and medical literatures demonstrate that investments in nutrition for pre-pubescent girls can improve health and accelerate sexual maturation (Garn 1987; Cooper et al. 1996; INSERM 2007). Moreover, nutritional shocks (both positive and negative) prior to puberty can have immediate and especially large effects for the onset of puberty and age of menarche for girls who had previously experienced early childhood undernutrition (Mul et al. 2002; Parent et al. 2003; Gluckman and Hanson 2006),[13,14] as was likely in the study population.[15] Therefore, in addition to affecting nutritional status directly (for example as measured by body mass index or

[13] Poor early life and later childhood nutrition and health are associated with delayed menarche, and better nutrition and health with early menarche. Medical research shows menarche is related to a minimum body fat mass, and that nutritional status in childhood affects body fat mass and menarche through the leptin hormone which helps regulate energy balance. Leptin levels fluctuate with nutritional intakes pointing to a mechanism by which positive (negative) nutrition shocks can translate into relatively immediate acceleration (delay) of menarche (Blum et al 1997).

[14] In a dynamic human capital production function starting in early childhood the relationship between early life deprivation and greater sensitivity to investment prior to menarche could be represented as a dynamic substitution. Because our conceptual framework (and data) begins in late childhood, we do not explicitly incorporate this complexity. What is most pertinent in our study, with girls having on average experienced poor childhood nutrition, is the potentially strong and immediate sensitivity of sexual maturation to nutritional shocks in early adolescence.

[15] Early childhood nutritional status for the sample is unavailable, but ~50% were stunted (height-for-age z-score < -2) as young women in 2010 and ~40% of girls < 3 years old in 2000 from the same localities were stunted.

BMI), nutritional investment patterns resulting from the CCT can affect the timing of female sexual maturation, with the interval just prior to puberty likely more sensitive to investment.

Median age of menarche in the sample was 13, pointing to period 2 as the more sensitive period to investment in health and the nutrition shocks for the following reasons. In 2000, a positive nutrition shock may have accelerated menarche in early treatment for some of the older girls in the 9–12 cohort but would have come too early to affect menarche for the majority of them. In 2003, the cessation of transfers for girls in the early treatment group (when they were 12–15 years old) likely meant a negative nutrition shock occurring during the more sensitive ages for the younger girls in the cohort, possibly delaying their menarche. Girls in the 9–12 cohort in late treatment, on the other hand, benefitted from the CCT starting in 2003 when they were 11–14 years old, including many 11–13-year-olds who had likely not reached menarche. This positive nutrition shock in the late treatment group at the beginning of period 2 could have accelerated the onset of menarche for them. Hence, for the girls in the 9–12 cohort overall, the start of the late treatment (period 2) was characterized by opposing nutrition investment shocks at a sensitive period: negative for early treatment, positive for late treatment. Given that the late treatment group benefited from the public investment in period 2, all else equal we expect earlier sexual maturity for late treatment girls (along with earlier fertility and marriage), and therefore negative estimated long-term differentials on these health-related outcomes. Based on the ages outlined above we also hypothesize negative effects would be concentrated among the youngest girls in the 9–12 cohort, 9–10-year-olds at baseline.

The above discussion outlines our expectations regarding sensitive periods for education (period 1) and health (period 2, for girls) investment. The conceptual framework (Section III.A), however, makes clear that final outcomes also reflect self-productivity and dynamic complementarities. For example, strong self-productivity and dynamic complementarities following early public investment in education in period 1 can increase education human capital for those in early treatment and the eventual observed differential between early and late treatment. There also can be later dynamic complementarities in period 2 related to possible synergies between education and health. At the same time, because it was known from the outset that the program would expand to the late treatment, the possibility of period 2 self-productivity and dynamic complementarities means there were incentives for households in late treatment to increase investment in period 1, before the program even began in their localities. This could lead to smaller long-term differentials relative to a context without such possibilities.

The investments in both education and health in periods 1 and 2 ultimately can affect labor market outcomes in period 3, and those effects can differ by sex. For both sexes greater program investment during education-sensitive period 1 could yield positive differentials in labor market outcomes in period 3, through positive differentials in schooling, learning and socio-emotional skills. For women, greater program investment in period 2 nutrition could lead to better nutritional status around the time of adolescent sexual maturation, resulting in earlier sexual maturity and possibly earlier fertility in period 3. Furthermore, the introduction of the adolescent health workshops and contraceptives in period 2 could have reinforced any possible physical sexual maturity-related behavioral changes including earlier sexual activity, potentially also leading to earlier fertility. Since both the nutritional and behavioral pathways described here are likely strongest in period 2, and we estimate program differentials for early minus late treatment, we would expect negative differential effects on fertility in period 3. Because higher fertility is negatively associated with most labor market outcomes for women, the possible fertility-channel related effects on labor market differentials could reinforce any positive education-channel effects on labor market differentials between early and late treatments.

IV. Data

A. Sources

Program Registry Census and Short-term Evaluation Surveys—A May 2000 program registry census done in all 42 localities provides baseline data on schooling, demographics and assets. Using the program census a random sample of 42 households in each of the 42 localities was drawn for the short-term evaluation; because locality sizes differed, we construct short-term evaluation sample weights to provide population estimates of the study area. A comprehensive household baseline survey was conducted in September 2000, with subsequent panel follow-ups in 2002 and 2004. Attrition was ~10% per round (Maluccio and Flores 2005).

2010 Long-term Evaluation Survey—The 2010 survey was conducted between November 2009 and November 2011, 9–11 years after the start of the program. The sample frame consisted of all households in the baseline 2000 evaluation sample plus, to increase power, a random oversample of households with children who were 11 years old in 2000 (specifically, born between January and June 1989). This age group was selected because they were at relatively high risk of dropping out of school in the years immediately following the start of the program, i.e., at ages when education might be particularly sensitive to investment (Section III.B). All long-term estimates account for the short-term evaluation survey sampling methodology and subsequent oversample using long-term evaluation survey sample weights constructed to provide population estimates of the study area. The main cohort of interest, children 9–12 years old at baseline in 2000 (Section III.B), has 2,200 individuals (1,138 males and 1,062 females).

The 2010 survey included separate household- and individual-level instruments. The household survey instrument collected both household- and individual-level information from the best-informed person available for the interview, generally the young adults themselves or the household head or spouse. It included individual schooling outcomes and detailed participation and earnings measures for all economic activities over the last 12 months. A complete labor market history including participation, location and earnings for all non-agricultural self-employment and wage jobs ever held was also collected. Location of activities was collected to distinguish between work performed while residing in the current home community versus work performed during periods of temporary migration. Given the mobility of this young population and temporary nature of many economic activities, such a comprehensive approach was key to characterizing labor market returns.

The individual survey instrument was administered to all individuals born after January 1, 1988. Data was collected through direct interviews of the young adults in their homes and designed to measure individual learning, cognition, socio-emotional, marriage and reproductive health outcomes. Weight and height were collected for females, but not for males.[16] To measure learning we implemented two math and three Spanish language achievement tests. The math tests included numerical fluency and a test of problem solving similar to the Peabody Individual Achievement Test (Markwardt 1989). Spanish tests included word identification, spelling and reading fluency. In addition, we administered two tests that capture both learning and cognitive development: the Spanish version of the Peabody Picture Vocabulary Test (PPVT) (Dunn et al. 1986) assessing receptive vocabulary and a forward and backward digit span test assessing memory. To measure cognition, we administered Raven's colored matrices (Raven et al. 1984).

[16] Anthropometry for males was not collected due to sensitivities related to potential military service in Nicaragua.

The Raven was included to help capture general non-verbal cognitive skills as opposed to the more specific skills likely acquired in the classroom. A strength of all the tests is they provide observed, as opposed to self-reported, measures of learning and cognition regardless of schooling status, substantially reducing concerns about social desirability or selectivity biases. Appendix D provides further detail on test administration protocols. We measured socio-emotional outcomes using the Center for Epidemiologic Studies of Depression (CESD) Scale and the Strengths and Difficulties Questionnaire (SDQ). Lastly, the individual-level instrument included a retrospective sexual and reproductive health module collecting information on prior attendance at and learning in the adolescent health education workshops, age of menarche for women, age of first sexual activity, and fertility and marriage. The module was developed based on qualitative interviews done in preparation for the 2010 survey that revealed perceptions of earlier sexual maturity and increased early sexual activity among teens during the program, the latter a possible unintended consequence related to provision of information and contraception (CIERUNIC 2009).

Nicaraguan Population Censuses—Finally, we also use the 1995 and 2005 Nicaraguan national censuses with individual-level residential, schooling, demographic and migration data.

B. Attrition, Internal Validity and Weighting

Due to extensive national and international tracking, attrition in the 2010 long-term evaluation survey is on par with or lower than in related longitudinal studies of similar populations and time horizons (Appendix E). Individuals not found at their original residences were tracked to new locations throughout Nicaragua and Costa Rica (the destination of 95% of the international migrants). Only 10% of men and 16% of women in the main cohort of interest, 9–12-year-olds at baseline, could not be tracked to their 2010 household location and are therefore missing the labor market outcomes and detailed schooling and marital status information collected in the household-level instrument. Attrition is higher, 19% for men and 22% for women, for the individual-level instrument which included the in-person tests. For individuals who could not be located we collected some schooling, demographic, migration and labor market information through proxy reports from other original household members, reducing the attrition on a set of basic indicators to 4% for men and 6% for women.

There are no significant differences in attrition (or in permanent migration) levels between early and late treatment; all differences are less than $|0.015|$ with p-values greater than 0.6.

Attrition also does not affect the balance of baseline observables (Appendix Table A1). However, attrition is correlated with baseline characteristics associated with migration, and those correlations differ between early and late treatment groups. Consequently, even with balanced attrition sample selectivity remains a potential concern for internal validity. Our preferred estimates account for attrition selection using inverse probability weights (IPW) constructed to allow for differences between early and late treatment groups and incorporating information from the survey tracking process. Individuals who were more difficult to find and interview are given higher weight since they were more similar to individuals not interviewed (Molina-Millán and Macours 2017). Separate IPW weights are calculated by sex for outcomes from household- and individual-level survey instruments, respectively, and all weights also incorporate the long-term evaluation survey sample weights to provide population estimates of the study area. Appendix E presents additional information on tracking, attrition and IPW construction.

C. 2010 Labor Market and Human Capital Families of Outcomes

We examine a wide range of labor market and human capital-related outcomes in 2010, organizing specific measures into different domains, or families of outcomes. Each family is constructed from the average of the z-scores of the individual components of that family.

To characterize the labor market for this predominantly rural, young adult population, we construct two families: labor market participation—capturing participation and temporary migration for work; and earnings—capturing labor market returns via earnings for work off the family farm. We consider two different versions of the earnings family to account for the skewed nature of the unconditional distribution of earnings in which many earn nothing (the small fraction not working as well as those only engaged in unremunerated economic activities) and an extended right-side tail: 1) ranks of earnings (Athey and Imbens 2017); and 2) reported earnings trimmed at the top 5% of values. Specific measures included are defined in results Section VI.B.

To explore various possible mechanisms underlying potential labor market effects, we expand upon the two forms of human capital outlined in the conceptual framework and construct schooling, learning, cognition, fertility and marriage, and socio-emotional families.

For schooling outcomes, the family includes whether the respondent was enrolled in school, highest completed grade and whether they had completed fourth grade, after which they were no longer eligible for the education transfer or subject to its conditionalities. To analyze learning and cognition outcomes we categorize the various tests into three families. The first, learning, focuses on the skills most likely acquired in the classroom and comprises the five achievement tests (math fluency, math problems, word identification, spelling and reading). The second, learning-and-cognition, includes the tests we expect are likely to capture both learning and cognition (receptive vocabulary and memory). The third, cognition, contains only the Raven.

For fertility and marriage outcomes, the components included are different for men and women. For both the family includes whether the individual had sex by age 15, had any children and had ever been married (including informal unions). For women, the family also includes age of menarche. The sign for age of menarche is reversed when averaging the z-scores for the family index so that higher values for the family are associated with higher fertility. We also consider an alternative version of the fertility and marriage family for women that includes BMI for non-pregnant women; unlike the other indicators in the family, BMI has the advantage of being objectively measured rather than self-reported. Higher BMI in this population can be indicative of better nutritional status, as well as earlier onset of puberty or prior childbearing.

For the socio-emotional outcomes, we conducted exploratory factor analysis combining all items from the CESD and SDQ (Appendix F).[17] The analysis reveals four latent factors we interpret as broadly capturing optimism, positive- and negative self-perception, and stress. The signs for negative self-perception and stress are reversed when averaging z-scores for the family index, so that higher values for the family indicate more positive socio-emotional outcomes.

V. Methodology

A. *Empirical Specification for 2010 Long-term Experimental Differentials*

We estimate intent-to-treat (ITT) effects of the randomized timing of exposure to the CCT using the following linear specification:[18]

$$Y_{il} = \alpha T_l + \boldsymbol{\beta X}_{il} + \varepsilon_{il}, \qquad (2)$$

[17] We use exploratory factor analysis because the correlations among items in the CESD and the SDQ suggested standardized scoring was unlikely to reflect the intended latent traits, similar to Laajaj and Macours (2021).

[18] Because the intervention and short-term evaluation began in 2000 and explicit plans for longer-term follow-up were not made until after the program had ended, there is no formal pre-analysis plan. Analyses adhere to the original experimental design and examine hypotheses outlined in proposals developed for the long-term follow-up.

where Y_{il} is the 2010 (i.e., period 3) outcome for individual i in baseline locality l. T_l is an ITT indicator equal to one for localities randomly assigned to early treatment and zero for late treatment. It measures the impact of early versus late exposure to the entire CCT package. Analyses use all respondents from both treatment groups in the age cohort being considered, regardless of initial completed schooling or actual program participation. All regressions also include strata fixed effects to account for randomization stratified by the locality-level marginality index. Given the randomized assignment, our main specification limits the other control variables in $\boldsymbol{X}_{il}$ to age of the individual at the start of the program in November 2000 using indicators for 3-month age groups, for whether they had 0, 1, 2, 3 or 4+ grades completed at baseline and regional fixed effects. Regressions are weighted to account for sampling and attrition providing population estimates of the study area. We assess robustness of the main specification to alternative controls, samples and approaches for addressing attrition selection in Section VI.B. The latter include estimates: 1) without IPW (using only the long-term evaluation survey sample weights with no correction for attrition); 2) with an alternative IPW that does not incorporate tracking information; 3) of Lee bounds; 4) of Kling and Liebman sensitivity bounds; and 5) of effects on a set of basic indicators using proxy reports for which attrition is lowest.

To reduce concerns related to multiple hypotheses testing, we organized individual variables into families as described above, and construct an index for each by averaging the z-scores of the specific components of the family (Kling et al. 2007). We assess robustness of this approach by also constructing alternative indices using inverse covariance weighting, which assigns less weight to individual indicators that are highly correlated within the families. Standard errors are adjusted for clustering at the locality level. We adjust for multiple hypotheses testing of all families at the same time using Anderson's (2008) familywise error rate. We also assess whether accidental imbalance related to the limited number of clusters (N=42) changes the significance, using randomization inference and estimating exact p-values under the sharp null hypothesis that the treatment effect is zero for all participants (Athey and Imbens 2017; Young 2019).

VI. Results

A. 2010 Long-term Experimental Differentials: Pooled Samples

To analyze the effects of the differential timing of exposure we estimate the 2010 long-term experimental differences between the early and late treatment groups, referred to below as the (long-term) differentials. Table 1 presents results for each family pooling men and women. For both the 7–12 (panel A) and 9–12 (panel B) cohorts there are significant positive long-term differentials in labor market participation (0.17–0.20 SD) and in earnings (0.12–0.14 SD). The results demonstrate that earlier exposure to the CCT translated into better labor market outcomes a decade later when individuals were beginning their adult working lives. The differentials in the schooling and learning families are approximately 0.10 SD for the 9–12 cohort, but more modest, 0.06 SD (and insignificant for learning), for the 7–12 cohort. Estimates are close to zero and insignificant for the socio-emotional family.

B. 2010 Long-term Experimental Differentials: 9–12 Cohort by Sex

To shed light on the possible sensitive periods and mechanisms described in the conceptual framework, we focus on the 9–12 cohort by sex (panels C and D). The pooled results mask differences in estimated magnitudes by sex, though only learning is significantly different between them (panel E).[19] Estimated labor market differentials are nearly twice as large for men. Both sexes have similar-sized positive differentials in schooling, but there is a significant learning differential only for men, with an estimated zero differential for women. In contrast, there is a significant negative differential in fertility and marriage for women, but none for men.

To unpack these results, the subsequent tables show the long-term differentials in the components of the outcome families, for each presenting first the men and then the women.

Labor market participation and earnings—As shown in Table 1 (panel C) and reproduced in column 1 of Table 2 (panel A), there is a sizeable 0.27 SD long-term differential in the labor market participation family for the male 9–12 cohort. Off-farm work is 6 percentage points higher (7% of the mean in late treatment) in early treatment. Men in early treatment are 9 percentage points more likely to have migrated temporarily for work in the last 12 months, nearly one-third higher. Possibly reflecting a migration for work strategy, they are also 8 percentage points (one-third) more likely to have ever had a salaried non-agricultural wage job, and 7 percentage points (one-half) more likely to have ever had an urban wage job.[20]

---

[19] Results for the 7–12 cohort by sex are qualitatively similar, but with smaller estimated magnitudes (Appendix Table A2). Results for the younger 7–8 cohort by sex are reported in Table 7 and discussed in Section VI.C.

[20] To provide further context for the findings, we note that nearly all men (98%) are working and most combine work on the family farm (89%) with work off the family farm (83%). In line with low average schooling (5.5 grades), most off-farm employment is unskilled. Men work as agricultural laborers on farms not belonging to the household or on large plantations, in salaried jobs in the non-agricultural sector (e.g., as construction workers or

In Table 3, we investigate whether differential labor market participation for men is accompanied by differential earnings.[21] To avoid selectivity bias all analyses are unconditional, with zero earnings for each component element when the individual did not report that type of earnings in the reference period. Results are broadly consistent across the various indicators, with an overall differential of about 0.2 SD for both rank (panel A) and 5% trim (panel B) versions of the earnings family. In addition, density functions in Appendix Figure A1 and quantile regressions in Appendix Table A3 demonstrate there are positive differentials across the earnings distribution. Because virtually all men (98%) are working, differential earnings do not stem from increased participation on the extensive margin, but rather from shifts in temporary migration for work and in the mix of economic activities.

Women in the 9–12 cohort exposed to the CCT program earlier also do markedly better in the labor market in 2010. There is a 0.17 SD differential in the labor market participation family (Table 2, panel B) driven by higher off-farm work (7 percentage points or 15%) and a doubling of temporary migration for work (9 percentage points). Correspondingly, there is a marginally significant differential in earnings of about 0.1 SD for both earnings families (Table 3, panels C and D). As was the case for the men, the findings suggest differentials in earnings for women stem from shifts in temporary migration for work and in the mix of economic activities.

Earlier exposure to the CCT led, after 10 years, to better labor market outcomes for both men and women. Results by sex are qualitatively similar but estimated magnitudes are larger for men than for women. One possible explanation for more modest differentials for the women is that less than half had earnings from off-farm labor and analyses are unconditional. Examination of

---

security guards), and in non-agricultural self-employment. In the poor, remote rural communities where the CCT operated, however, opportunities for even these types of low-skilled employment are often limited. For this reason, we incorporated a measure of seasonal or temporary migration for work in the labor market participation family.

[21] We use several measures to capture earnings from different activities (agricultural or non-agricultural and salaried or self-employment) that are often seasonal or temporary (Table 3). We do not, however, include earnings from the family farm since person-specific individual returns could not be reliably quantified in this context. Only 11% of the cohort of young men are head of their own household so that nearly all work on the farm of an older household member, typically their father or father-in-law. Moreover, we did not collect detailed information on agricultural inputs, precluding calculation of earnings for household farms. Although in theory the positive differentials for off-farm earnings we estimate could be offset by lower on-farm earnings, patterns of participation make it unlikely. First, the estimated differential in participation in on-farm work is small and positive, indicating that higher off-farm work did not substitute for on-farm work on the extensive margin. Second, there is no differential in the total number of months worked off-farm; average months working off-farm during temporary migration increases while average months working off-farm (while residing in the current home community) decreases. These patterns suggest minimal scope for crowding out on-farm work on the intensive margin.

the human capital-related outcomes below suggests that sex differences in other mechanisms also may help explain sex differences in labor market outcomes for the young adults.

Schooling, learning and cognition—There is a ~0.1 SD long-term differential in the schooling family for men (Table 4, panel A), including nearly 0.3 more grades completed (5%). Early treatment men are also 5 percentage points (25%) more likely to still be in school. Overall, a substantial minority (18%) in the late treatment group are still studying in 2010 when they were 19–22 years old, with more than 80% of them in secondary school where they typically enroll in weekend programs that enable them also to work. Average completed grades for early treatment men remain low (5.8), however, and only 78% have completed fourth grade.

Higher schooling for early treatment men is accompanied by higher learning (0.18 SD). In particular, there are significant long-term differentials in exactly the sorts of skills taught in school: 0.16 SD for math and 0.20 SD for Spanish. The findings for Spanish are corroborated by a significant differential on self-reported literacy (reported in column 5 though not included in the learning family which comprises only direct tests). Magnitudes of the differentials are sizable and in line with absolute impacts of education interventions in other similar settings (Evans and Yuan 2022). Together, the results for schooling and learning demonstrate that men exposed to the CCT earlier not only have higher completed grades and do better on achievement tests, but also are more likely to still be studying, suggesting observed positive education differentials may have continued to grow even after 2010.

Differentials are smaller and insignificant for the mixed learning-and-cognition family that arguably captures both classroom and non-classroom skills. The estimate for the Raven, a cognitive test less likely to reflect skills directly acquired in the classroom, is close to zero (-0.02 SD). The lack of differentials for cognition-related measures is as expected given that the intervention for the 9–12 cohort began in late childhood, well after what are often regarded as the most sensitive ages for investment in cognitive development.[22]

For women there is a positive long-term differential in the schooling family the same size as for men, about 0.1 SD (panel B). The differential between early and late treatment in grades completed is more modest (0.18 or 3%) and not significant (though there is a significant

[22] It is for this reason that the cognition-related families are not included in Table 1. In other work we examine effects on these measures for cohorts exposed to the CCT at younger ages (Barham et al. 2013). The distinction between achievement and cognitive tests was outlined in proposals developed for the 2010 long-term evaluation.

differential in having completed fourth grade), possibly related to lower potential sensitivity to education investment for females in period 1 due to later typical ages of dropout. Indeed, even in 2010, 30% of the young women are still in school, with two-thirds in secondary school and one-third studying at the tertiary level.

In contrast to the men, there are no significant differentials in learning for women. Point estimates are close to zero for the various families (panel B), as well as for the underlying individual tests (Appendix Table A4) and for self-reported literacy (panel B). The lack of a differential could indicate that the CCT did not improve learning for women in either treatment group but is also consistent with positive similarly sized absolute program effects for each treatment group. We provide short-term experimental and medium-term non-experimental evidence supporting the latter possibility in Sections VI.D and VI.E below.

The long-term differentials in schooling and learning are consistent with the hypothesis that period 1 was more sensitive for education investment in the 9–12 cohort, especially for boys.

Fertility and marriage—There is no overall significant differential in the fertility and marriage family for men (Table 5, panel A).[23] Individual components offset one another with a positive differential for ever having sex before age 15, but a negative one for ever married.

For women on the other hand (panel B), there is a significant negative differential of 0.17 SD for both the family with and without BMI (columns 1 and 2). There is, notably, a significant differential in the age of menarche, with the late treatment group having reached sexual maturity three months before the early treatment group. Correspondingly, there is a large negative differential on measured BMI, 0.7 kg/m^2 lower in early treatment.[24] The early treatment group is 11 percentage points less likely to have had sex by the time they were 15 years old, one-third lower than the average in the late treatment. This large differential is mirrored by a 6-percentage point (10%) lower probability of having a child by 2010 (p-value 0.17).[25] Overall the results for the fertility and marriage family for women are in line with greater period 2 sensitivity to

---

[23] Because anthropometrics were not collected for men, we are unable to evaluate nutritional mechanisms related to health human capital for men to the same extent as for women.

[24] It is possible that recall error affects reports of age of menarche, though there is no obvious reason such measurement error would affect the early and late treatment groups differently. Results on BMI for non-pregnant women, objectively measured and therefore unlikely to suffer from differential measurement error, are significant and in line with the other reproductive health outcomes including age of menarche. The consistency in findings across outcomes suggest systematic recall measurement error on age of menarche is not driving the fertility results.

[25] Analysis of the short-term evaluation data provides evidence consistent with this, as an estimated negative 3-percentage point differential in fertility had already emerged by 2004 when the girls were still only 13–16 years old.

investment and the nutrition shocks (resulting from the end of the program in early treatment and the start in the late treatment) operating through the age of menarche. In Section V.E we discuss non-experimental evidence that this negative differential in fertility does not appear to be driven by an absolute increase in fertility in late treatment.

As outlined in Section III.B, given the median age of menarche, we expect the period 2 sensitivity to be larger for the younger girls in the cohort and therefore examine the fertility and marriage family and its components separately for girls 9–10 and 11–12 years old at baseline in Appendix Table A5. The differential for the family for the 9–10 cohort is indeed larger (-0.24 SD, significant at 1%) than for the 11–12 cohort (-0.10 SD, insignificant). Moreover, although not statistically different, larger differentials are observed for the younger cohort in all underlying components; point estimates for age of menarche indicate the younger girls in the late treatment group reached sexual maturity four months before the early treatment, compared to two months earlier for the older girls. This pattern of results for age of menarche and fertility further supports the possibility of greater sensitivity to nutritional investment and its effects on sexual maturation in period 2 and suggests nutrition is playing an important role. In Section VI.D we use the short-term evaluation surveys to document large positive impacts on food availability and quality providing evidence supporting such program-induced nutritional gains.

In columns 8–9 of Table 5 we examine outcomes related to the adolescent health workshops for the 9–12 cohort. As expected, they demonstrate that early treatment girls were less likely to have attended the workshops which began in 2003, corroborated by their being less likely to know about a key health topic emphasized in them, the Pap smear test. If the workshops (or contraceptive provision) had the unintended positive effects on sexual activity suggested by the qualitative research, differential exposure to them could have exacerbated potential negative differentials on fertility induced by the nutritional shocks and physical maturity described above.

To investigate this possibility, we consider an alternative comparison, contrasting girls 9–12 years old in 2000 (our main 9–12 cohort) in early treatment with a younger cohort of girls 9–12 years old in 2003 in late treatment. This provides a partial test of whether differences (including the workshops or the modestly lower transfer amounts) between the early and late treatment programs drive the differential outcomes in 2010. Appendix Table A6 demonstrates there are only small and insignificant differences between the two groups for age of menarche and whether they had had sex by age 15.[26] Moreover, there are no differences in having completed fourth grade for either sex. These results are consistent with the 2010 long-term differentials in fertility coming from the timing of the nutrition shocks during a sensitive period for sexual maturation in period 2, rather than from any differences in the structure of the program between early and late treatment. Nevertheless, with the available data we cannot rule out the possibility of some differential behavioral effects related to the workshops at older ages.

Socio-emotional Outcomes—Socio-emotional outcomes represent another potential important pathway through which the program could have influenced long-term labor market outcomes. The teenage years have been hypothesized to be a period sensitive to investment in socio-emotional (or non-cognitive) skill formation (Cunha and Heckman 2007), so that differences in timing of program exposure could well lead to differentials in such skills. Interpretation as a possible mechanism for the labor market results is arguably less straightforward than for education or fertility, however, since it is plausible that the causality runs in the opposite direction as well, with labor market outcomes affecting current socio-emotional measures.

Table 6 shows small and insignificant long-term differentials in the socio-emotional outcome families for men and women. Examining the each of the factors separately, however, reveals that for men the average masks offsetting differentials in the various latent traits (panel A). Specifically, men in the early treatment group are more optimistic and have a more positive self-perception (~0.25 SD), possibly reflecting their higher learning and earnings. Meanwhile, early treatment men also are more likely to exhibit greater stress and agree with statements reflecting negative self-perception, although point estimates for these traits are smaller (~0.16 SD). This pattern of offsetting effects echoes Fernald et al. (2008) who find that increased economic opportunities improved some aspects of mental health but at the same time increased stress. For women (panel B), there is also a significant differential for negative self-perception. Overall, the findings do not provide clear evidence that the labor market differentials result from, or for that matter drive, differentials in socio-emotional outcomes for either sex.

Robustness—The main results for the 9–12 cohort by sex for the different families of outcomes are robust to adjustments for multiple hypotheses testing (using Anderson's familywise error rate) and randomization inference (Table 1, panels C and D). For the latter, we

[26] Because this is a comparison of girls who across treatment groups have different ages when measured in 2010, we do not compare age dependent outcomes such as ever married or number of children.

can reject the sharp null hypothesis that all treatment effects are zero for all the outcome families by sex that were significant using conventional standard errors clustered at the locality level. Appendix B also demonstrates that the results are robust to: 1) alternative weights and samples; 2) different assumptions related to attrition, confirming that the main findings are not driven by IPW reweighting; and 3) a family index constructed using inverse covariance weighting.

Finally, the logic that timing of early versus late treatment can affect some age groups differently means it may be possible to learn more about the mechanisms by examining differentials for narrower age cohorts, as done for the female fertility and marriage family above. This strategy comes at the cost of reduced power, particularly for younger cohorts for whom there was no targeted oversample. We present the 2010 long-term experimental differentials for the two-year age cohorts in Table 7 (which also includes the younger 7–8 cohort discussed below in Section VI.C). For men, differentials are similar across the 9–10 and 11–12 cohorts, with point estimates a little larger for most outcomes for the 11–12 cohort including a positive significant differential in the socio-emotional family. Results for women are broadly similar across the two-year cohorts, apart from fertility discussed above and a now significant negative differential in the socio-emotional family for the older 11–12 cohort. The differentials in the socio-emotional families for the 11–12 cohorts by sex are in the same direction as the corresponding differentials for the learning families of this cohort, consistent with learning and socio-emotional skills being sensitive in the same periods. The pattern is also observed for the 9–12 cohort (Table 1, panel E).

C. 2010 Long-term Experimental Differentials: 7–8 Cohort by Sex

Based on the age-eligibility cut-offs, children in the 7–8 cohort (i.e., 7–8 years old in 2000) were potentially eligible for the CCT education component in either treatment group throughout program operation (Appendix Table C1). The logic used in Section III.B for identifying education-sensitive periods for investment therefore points to period 2 for this cohort (when they were reaching ages at which school dropout risk rises), although it is less definitive. With less contrast in the potential direct exposure to the education component of the program across the two experimental groups, the differential could depend relatively more on whether possible gains from earlier public investment, or later positive self-productivity and dynamic complementarities in early treatment, dominate gains from any anticipation effects combined with public investment in period 2 in late treatment. Examining investment in health, period 2 is also arguably more sensitive, with most females in the 7–8 cohort too young to reach puberty in period 1 but 12–13 years old by the final year of transfers in late treatment. Because the program ended in 2005, however, it also may be that there would not have been sufficient exposure at ages close to 13 to have had a substantial influence on sexual maturation.

We present the 2010 long-term experimental differentials for the families for the 7–8 cohort by sex in Table 7 (panels A and D). Differentials for both males and females are relatively small, and all are statistically insignificant. The lack of differentials is likely a result of this cohort being exposed to the full three years of the two core program components in both the early and in the late treatment groups as described above. It does not necessarily mean there were no positive absolute effects on these cohorts. We provide short-term experimental and medium-term non-experimental evidence examining this possibility below in Sections VI.D and VI.E.

D. Short-term Experimental Impacts on Schooling and Nutritional Investments in Periods 1 & 2

To help unpack the dynamics underlying the 2010 long-term experimental differentials, we use the short-term evaluation surveys to document program compliance and estimate short-term experimental impacts on schooling and household food availability in 2002 and 2004, corresponding roughly to periods 1 and 2 in the conceptual framework.

School enrollment during the program—Enrollment patterns during program implementation demonstrate that compliance in both treatment groups was high and in line with the timing of the public investment and application of the education component age criteria and conditionalities. As shown in Table 8, in early treatment localities in 2002 (period 1) enrollment was nearly universal for the 7–12 cohort. In late treatment in 2004 (period 2), after the program had been operating there for almost two years, enrollment was similarly close to 100% for the 7–8 cohort, suggesting high compliance for that age-eligible cohort. Enrollment was lower in late treatment for the 9–12 cohort, most of whom had aged out by 2004 (when 13–16 years old).

Short-term experimental effects on schooling—To examine the short-term effects of the public investments during program implementation, we use equation (2) and estimate the experimental differential effects for the 9–12 cohort by sex in 2002 and 2004. In 2002—when early treatment localities had received transfers for two years but late treatment localities had not yet benefited—boys in the 9–12 cohort have 0.36 additional grades (15% more) and are more likely to have completed fourth grade (Table 9, panel A). Moreover, they are 18 percentage points more likely to be enrolled (25%), 36 percentage points more likely to be regularly attending (66%) and 15 percentage points more likely to be literate (20%). Estimated effects in 2002 for girls (panel C) are qualitatively similar and also significant, but smaller. The larger effects for boys in the short-term are consistent with period 1 having been more sensitive to investment in education for them, as hypothesized.

The short-term results for 2002, rather than simply measuring the absolute effect of early treatment compared to a pure experimental control, also could be influenced by anticipatory behavior in the late treatment. In the conceptual framework we argued that with positive dynamic complementarities advance notice of future public investment in period 2 could incentivize households in late treatment to increase investment in period 1. This could lead to smaller period 1 differentials relative to a pure control group. In contrast, however, because of the program eligibility rules it is also conceivable that late treatment households would not enroll or would hold back children so that they would not surpass fourth grade before period 2, particularly if dynamic complementarities are weak. This could lead to larger period 1 differentials. Enrollment and repetition rates for the 9–12 cohort in late treatment show that, if anything, higher earlier investment was more likely. Enrollment for late treatment children was 77% at baseline in 2000 and reached 80% by 2002, increasing for both sexes but with larger increases for boys. It also does not appear that children were systematically being held back to avoid completing fourth grade (thus remaining eligible for transfers longer); repetition rates for the same cohort were 12% in 2000 and 10% in 2002. Therefore, positive differentials in 2002 likely underestimate the absolute program effects.

Between 2002 and 2004 the differential in grades completed (as well as the fraction completing fourth grade) continues to widen, reaching 0.49 grades for boys (panel B) and 0.57 for girls (panel D). Alongside these increases, however, and in line with the program having ended in the early but begun in the late treatment, effects on enrollment and regular attendance in 2004 are reversed. The 2004 negative differentials between early and late treatment for enrollment (14 percentage points or 18%) is significant for girls, suggesting that late treatment children were already beginning to catch up on schooling. Notably, this occurs even though most girls in the late treatment in the 9–12 cohort had already aged out and were no longer themselves directly eligible for the education component or subject to its conditions. The negative differential on enrollment for girls in 2004, therefore, likely stems from other program components that households of late treatment girls were receiving or is related to dynamic production function technology such as synergies between education and health human capital production. These short-term patterns are consistent with expectations about effects from the public investment and point to the potential for considerable catch-up in grades completed for the late treatment girls, indicating that for them more so than for the boys, long-term differentials on schooling could underestimate absolute effects. This is corroborated by the 2010 survey showing that both the early and late treatment groups continued to advance in school after 2004 (when the average was 4.4) reaching 6.8 grades completed for the late treatment group in 2010 (Table 4).

Short-term experimental effects on nutritional investments—Turning to the effects of the public investment on nutrition-related indicators, in Table 10 we use short-term evaluation data to confirm that exposure to the CCT is related to large experimentally induced shocks in the availability of nutritious food in periods 1 and 2 for the cohorts of interest. Both the quantity and the quality of household food consumption changes significantly. In 2002, per capita food consumption expenditure in households with a boy (panel A) or with a girl (panel C) in the 9–12 cohort is ~35% higher in early treatment. Moreover, there is improved nutritional quality of food with higher consumption shares of animal proteins and fruits and vegetables, alongside lower shares of staples or other less nutritious foods.[27] By 2004 (period 2), when the CCT had been operating for more than a year in late treatment but had been phased out in early treatment, the patterns reverse, with 13% or more higher food consumption in late treatment.[28]

E. Non-experimental Differences-in-differences (DD) Absolute Effects

As discussed above, the absence of long-term differentials for some measures (grades completed and learning for the female 9–12 cohort) or some cohorts (both the male and female 7–8 cohort) could result from similar sized absolute effects that cancel one another out in the estimated differential. The short-term experimental evidence presented above (and for the 7–8 cohort in Appendix Table A7) points to this possibility for certain outcomes. In 2002 before the late treatment had begun to receive the program there were positive effects on schooling and

[27] Short-term improvements in quantity and quality of household food consumption are consistent with findings indicating significant reductions in stunting among children in the first years of life (Maluccio and Flores 2005).

[28] Absolute values of the differential effects are smaller in 2004 than in 2002, possibly because transfer sizes were modestly smaller but also because the early treatment group may have continued to invest in better nutrition even after the end of the program, as observed for a related Nicaraguan CCT program (Macours et al. 2012).

learning (measured by literacy) for boys and girls in both age cohorts, but by 2004, after it had begun receiving the program, the late treatment showed signs of starting to catch up.

To help further interpret what underlies the differentials, we explore (less well-identified) non-experimental differences-in-differences (DD) estimates of 5-year absolute effects comparing treatment municipalities to adjacent similarly poor non-treatment municipalities using 1995 and 2005 Nicaraguan population census data. Results point to large absolute increases in schooling outcomes after five years for both boys and girls in the 7–8 and 9–12 cohorts, with effects on grades completed twice as large as the experimental differentials (Appendix Table G1).[29] Together with the 2010 differentials, the findings suggest that late treatment children were largely able to catch up on schooling and learning with the early treatment, possibly as a result of program exposure in late treatment through 2005. Consequently, the experimental differentials may represent lower bounds for the absolute program effects.

Using the national census data and the same DD specification, we also examine early fertility for the 9–12 cohort of girls (14–17 years old at the time of the second census in 2005). Results indicate a 2-percentage point reduction in both marriage and fertility, equivalent to a 25% reduction for early fertility. Finally, in Appendix G we compare the 9–12 cohorts from the study with a same-age cohort of women from the 2011–12 Nicaraguan Demographic and Health Survey. The non-experimental comparison suggests the CCT is associated with an overall increase in BMI alongside a reduction in early fertility for the entire cohort (early and late treatment groups combined), making it unlikely that the negative experimental differential reflects an absolute increase in fertility in the late treatment group.

F. Interpretation and Implications for Long-term Labor Market Results

Although the randomized design does not allow fully disentangling the various mechanisms potentially underlying period 3 labor market outcomes, the sex-specific results for schooling, learning and fertility suggest those mechanisms differ for men and women.

For men, period 1 exposure to the CCT led to more schooling and learning, along with changes in labor market participation and higher earnings. The pattern supports a principal

[29] Appendix G provides the detailed methodology and results for the DD, demonstrates that results are robust to alternative comparison groups, and provides falsification test evidence in support of the common trends assumption.

rationale for CCTs and provides evidence of their potential to reduce the intergenerational transmission of poverty through improved (education) human capital.

Higher earnings for the men, however, were not the result of transformative increases in their schooling or changes toward skilled or formal employment. This is unsurprising considering the short-term nature of the program and that the improvements in basic skills were for a population with initially low levels of education. Findings from qualitative work suggest basic skills attained might have led to higher labor market returns via seasonal or temporary migration. Although such migration has associated costs, there are large wage differentials across regions in Nicaragua, and wages are substantially higher in Costa Rica. Improved math skills may have helped the young men assess the cost-benefit trade-offs of temporary migration, and improved reading comprehension skills may have enabled them to more easily complete required paperwork, particularly relevant for international migration to Costa Rica (CIERUNIC 2009).

For women, period 1 exposure also led to positive differentials in schooling, driven by completing the program-targeted fourth grade, but no significant differentials in total grades completed or learning. The absence of differentials in grades completed or learning is consistent with positive absolute program effects for both the early and late treatment groups (Section V.E).

The negative differential in the fertility and marriage family suggests that program-related nutrition shocks in period 2, along with modest positive differential in schooling described above, may have both contributed to the differential labor market outcomes for women in period 3. A fertility-related mechanism is consistent with the challenges Nicaraguan women face when combining motherhood and the responsibilities of marriage with labor market activities in a context where participation often requires commuting to nearby urban centers or migrating to larger cities or to Costa Rica.

G. Cost-benefit assessment

The positive differential effects on earnings suggest the potential for recouping the costs of the CCT. A full cost-benefit analysis would require monetary valuation of all resource costs and all potential absolute benefits of the program for each individual household member, including those not in the study, and is beyond our scope. The clearest evidence of positive returns is for men in the 9–12 cohort. Even for that group, however, differences in earnings between those with early versus late exposure likely understate longer-term annual returns because: 1) they only capture the differential earnings effects; 2) they do not reflect final schooling differentials which may have continued to widen, further increasing earnings differentials; and 3) they are measured at the beginning of the beneficiaries' working lives and therefore have potential to increase over time. With these caveats, using only the earnings gains for this cohort and under conservative assumptions regarding cost and benefit flows, the net present value with a 10% discount rate turns positive within 2–3 decades (Appendix I).

VII. Conclusions

Within the framework of a dynamic human capital production function, in this paper we provide experimental evidence that earlier exposure to a CCT in rural Nicaragua translated into sustained positive differences in labor market participation and earnings a decade after the program began, for both men and women. Results suggest, however, that the mechanisms through which the CCT program components generated the positive differentials differed by sex.

For males, exposure to the CCT during primary school ages when they were vulnerable to school dropout led to substantial education and learning gains compared to exposure a few years later. Subsequent employment gains were made possible by higher levels of temporary migration. A plausible interpretation is that with more education and learning the men developed core competencies enabling them to find higher paid work away from home when they were beginning their adult working lives, possibly setting them on a higher future earnings trajectory.

Results on sustained learning for the young men contrast with some findings in the literature that point to limited effects of CCTs on learning, findings that led a recent high-level panel to conclude CCTs may not be cost-effective for addressing learning gaps (World Bank 2020). Positive learning and income differentials in Nicaragua were substantial. At the same time, given low initial education levels and other remaining constraints, the program did not fully erase learning gaps or wholly transform the lives of the young men. Nevertheless, cost-benefit analysis suggest the earning gains were sufficient to recover investments within a few decades.

For females, earlier exposure to the CCT also led to positive differentials in schooling attainment, but not in learning. In addition, the timing of program-related nutrition shocks led to later onset of menarche and sexual activity, and lower BMI in the early compared to the late treatment group. These differences in reproductive health-related outcomes help explain the positive labor market differentials for the young women. For those in late treatment, the effect of the nutrition shocks on sexual maturation may have partly offset possible delayed fertility resulting from higher schooling, while for those in early treatment the opposite shock may have complemented any such schooling effects. The results underscore the importance of accounting for the nutrition and health implications—and not just education—of cash transfer or related programs when analyzing longer-term effects. In addition, the results suggest that both timing and program design are important; transfers and related nutritional shocks during preteen years can affect the age of menarche, so the specific ages at which interventions target girls can affect subsequent marriage, fertility and labor market outcomes.

Taken together, the evidence suggests that the specific timing of nutrition, health and education interventions during primary school ages can have substantial consequences for their effectiveness and long-term returns. Moreover, at these ages different types of human capital can have similar or reinforcing effects on subsequent outcomes but can also potentially affect important outcomes such as fertility and labor market participation in opposite directions. Future research should aim to further disentangle the roles of nutrition and education during these sensitive years, and their interactions, to ensure optimal returns.

VIII. References

Anderson, M. 2008. "Multiple Inference and Gender Difference in the Effects of Early Interventions: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects." *Journal of American Statistical Association* 103(484):1481–95.

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.

Araujo, M. C., and K. Macours, 2021. "Education, income and mobility: Experimental impacts of childhood exposure to *PROGRESA* after 20 years," *PSE working paper* 2021/10.

Athey, S., and G.W. Imbens. 2017. "The Econometrics of Randomized Experiments." Banerjee, A. and E. Duflo, (eds.), *Handbook of Economic Field Experiments*. Volume 1. Elsevier.

Attanasio, O., C. Meghir and E. Nix. 2020. "Human Capital Development and Parental Investment in India." *Review of Economic Studies* 87: 2511–41.

Avitabile, C, 2012. "Spillover Effects in Healthcare Programs: Evidence on Social Norms and Information Sharing" *IDB Working paper* 380.

Baird, S., J.H. Hicks, M. Kremer, and E. Miguel. 2016. "Worms at Work: Long-run Impacts of a Child Health Investment." *Quarterly Journal of Economics* 131 (4): 1637–80.

Baird, S., C. McIntosh, and B. Özler. 2011. "Cash or Condition? Evidence from a Cash Transfer Experiment." *The Quarterly Journal of Economics* 126(4): 1709–53.

Baird, S., C. McIntosh, and B. Özler. 2019. "When the Money Runs Out: Do Cash Transfers have Sustained on Human Capital Accumulation?" *Journal of Development Economics* 140: 169–85.

Bandiera, O., N. Buehren, R. Burgess, M. Goldstein, S. Gulesci, I. Rasul, and M. Sulaiman. 2020 "Women's Empowerment in Action: Evidence from a Randomized Control Trial in Africa", *American Economic Journal: Applied Economics*, 12(1):210–59.

Bandiera, O. R. Burgess, S. Gulesci, N.Das, I. Rasul, M. Sulaiman. 2017. "Labor Markets and Poverty in Village Economies." *Quarterly Journal of Economics*, 132(2): 811–70.

Banerjee, A., E. Duflo, R. Chattopadhyay, and J. Shapiro. 2016. "The Long Term Impacts of a "Graduation" Program: Evidence from West Bengal." Unpublished, MIT.

Barham, T. R. Kuhn, and P. Turner. 2018. "No Place Like Home: Long-Run Impacts of Early Child Health and Family Planning on Economic and Migration Outcomes." CU Boulder.

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): 467–71.

Barrera-Osorio, F., Linden, L.L., and Saavedra, J.E. (2019). "Medium- and Long-Term Educational Consequences of Alternative Conditional Cash Transfer Designs: Experimental Evidence from Colombia." *American Economic Journal: Applied Economics,* 11(3): 54–91.

Bastagli, F., J. Hagen-Zanker, L.Harman, V. Barca, G. Sturge and T. Schmidt. 2016. Cash Transfers: What does the evidence say? Overseas Development Institute, London, UK.

Behrman, J. R., S.W. Parker and P.E. Todd. 2009a. "Medium-Term Impacts of the *Oportunidades* Conditional Cash Transfer Program on Rural Youth in Mexico." In: S. Klasen and F. Nowak-Lehmann, eds. *Poverty, Inequality, and Policy in Latin America*, 219–70. Cambridge, United States: MIT Press.

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

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

Blum WF, Englaro P, Hanitsch S, A. Juul, N.T. Hertel, J. Muller, N.E. Skakkebaek, M.L. Heiman, M. Birkett, A.M. Attanasio, W. Kiess, and W. Rascher. 1997. "Plasma leptin levels in healthy children and adolescents: dependence on body mass index, body fat mass, gender, pubertal stage, and testosterone." *J Clin Endocrinol Metab*;82(9): 2904–10

Buchmann, N., E. Field, R. Glennerster, S. Nazneen, S. Pimkina, and I. Sen, 2018. "Power vs Money: Alternative Approaches to Reducing Child Marriage in Bangladesh, a Randomized Control Trial", mimeo, Duke University.

Cahyadi, N., R. Hanna, B.A. Olken, R.A. Prima, E. Satriawan, and E. Syamsulhakim. 2020. "Cumulative Impacts of Conditional Cash Transfer Programs: Experimental Evidence from Indonesia." *American Economic Journal: Economic Policy*, 12(4): 88–110.

Carneiro, P., I. Lopez Garcia, K.G. Salvanes and E. Tominey, 2021. "Intergenerational Mobility and the Timing of Parental Income", *Journal of Political Economy*, 129(3) 757–88.

Centro de Investigación y Estudios Rurales y Urbanos de Nicaragua (CIERUNIC). 2009. "Qualitative Findings for the Evaluation of the Long-term Impact of the *Red de Protección Social* in Nicaragua." Unpublished.

Cooper, C., D. Kuh, P. Egger, M. Wadsworth and D. Barker, 1996. "Childhood growth and age of menarche." *Br. J. Obstet. Gynaecol*. 103(8): 814–7.

Cunha, F. and J. Heckman. 2007. "The Technology of Skill Formation." *American Economic Review Papers and Proceedings* 97(2): 31–47.

Cunha, F., J. Heckman and S.M. Schennach. 2010. "Estimating the Technology of Cognitive and Noncognitive Skill Formation." *Econometrica* 78(3): 883–931.

de Janvry, A., and E. Sadoulet. 2006. "Making Conditional Cash Transfer Programs More Efficient: Designing for Maximum Effect of the Conditionality." *World Bank Economic Review* 20(1): 1–29.

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

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

Dunn, Lloyd M., D.E. Lugo, E.R. Padilla, and Leota M. Dunn. 1986. *Test de Vocabulario en Imágenes Peabody*. Circle Pines, Minnesota: American Guidance Service, Inc.

Duque, V., Rosales, M. F., and Sanchez, F., 2018. "How Do Early-Life Shocks Interact with Subsequent Human-Capital Investments? Evidence from Administrative Data." Mimeo, University of Michigan.

Evans, D., and F. Yuan, 2022. "How Big Are Effect Sizes in International Education Studies?", *Educational Evaluation and Policy Analysis*, 44(3):532-540.

Fernald, L., 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* 371:828–37.

Fernald, L., R. Hamad, D. Karlan, E. Ozer and J. Zinman. 2008. "Small individual loans and mental health: a randomized controlled trial among South African adults", *BMC Public Health* 8: 409.

Field, E., and A. Ambrus. 2008. “Early Marriage, Age of Menarche, and Female Schooling Attainment in Bangladesh.” *Journal of Political Economy* 116(5): 881–930.
Filmer, D., and N. Schady. 2014. “The Medium-Term Effects of Scholarships in a Low-Income Country.” *Journal of Human Resources* 49(3): 663–94.
Fiszbein, A., and N. Schady. 2009. “Conditional Cash Transfers: Reducing Present and Future Poverty.” *World Bank Policy Research Report*. Washington, DC: World Bank.
Garn, S.M., 1987. “The secular trend in size and maturational timing and its implications for nutritional assessments: a critical review.” *Journal of Nutrition* 117, 817–23.
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., 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.
Gluckman, PD and M.A. Hanson. 2006, “Evolution, Development and Timing of Puberty” *Trends in endocrinology and metabolism* 17(1): 7–12.
Khana, Madhulika. 2020. “The Precocious Period: Impact of Early Menarche on Schooling in India.” Unpublished manuscript, Georgetown Unitervisty.
Handa, S. L. Natali, D. Seidenfeld, G. Tembo and B. Davis, 2018. “Can Unconditional Cash Transfers Raise Long-term Living Standards? Evidence from Zambia”, *Journal of Development Economics*, 133: 42-65.
Haushofer, J. and J. Shapiro. 2018. “The Long-Term Impact of Unconditional Cash Transfers: Experimental Evidence from Kenya.” Unpublished, Princeton.
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): 411–16.
INSERM (Institut National de la Sante et de la Recherche Medicale), 2007. *Growth and puberty: secular trends, environmental and genetic factors*. A collective expert report.
Khan, M.E., A. Hazra, A. Kant and M. Ali, 2016. “Conditional and Unconditional Cash Transfers to Improve Use of Contraception in Low and Middle Income Countries: A Systematic Review”, *Studies in Family Planning*, 47(4): 371-383.
Kling, J., J. Liebman and L. Katz. 2007. “Experimental Analysis of Neighborhood Effects.” *Econometrica* 75(1): 83–119.
Laajaj, R. and K. Macours. 2021. “Measuring Skills in Developing Countries.” *Journal of Human Resources* 56(4): 1254–95.
Lamadrid-Figueroa, H. G. Angeles, T. Mroz, J. Urquieta-Salomon, B. Hernandez-Prado, A. Cruz-Valdez and M.M. Tellez-Rojo et al 2010. “Heterogeneous Impact of the Social Programme Oportunidades on Use of Contraceptive Methods by Young Adult Women Living in Rural Areas”, *Journal of Development Effectiveness*, 2(1): 74-86.
Macours, K., N. Schady and R. Vakis. 2012. “Cash Transfers, Behavioral Changes, and Cognitive Development in Early Childhood: Evidence from a Randomized Experiment.” *American Economic Journal: Applied Economics* (4)2: 247–73.
Macours, K. and R. Vakis. 2016. “Sustaining Impacts When Transfers End: Women Leaders, Aspirations, and Investment in Children.” *NBER Working Paper* No. 22871.
Maluccio, J. A., and R. Flores. 2005. “Impact Evaluation of a Conditional Cash Transfer Program: The Nicaraguan *Red de Protección Social.*” *Research Report 141*. Washington, DC, United States: International Food Policy Research Institute.

Markwardt, F. C. 1989. *Peabody Individual Achievement Test-Revised Manual*. Circle Pines, Minnesota, United States: American Guidance Service.
Molina-Millán, T., T. Barham, K. Macours, J.A. Maluccio, and M. Stampini. 2019. "Long-term Impacts of Conditional Cash Transfers: Review of the Evidence." *World Bank Research Observer*, 3(1): 34(1): 119–59.
Molina-Millán, T. and K. Macours. 2017. "Attrition in Randomized Control Trials: Using Tracking Information to Correct Bias." *CEPR Discussion Paper No. DP11962*
Molina-Millán, K. Macours, J.A. Maluccio and L. Tejerina. 2020. "Experimental long-term impacts of early childhood and school-age exposure to a Conditional Cash Transfer Program." *Journal of Development Economics* 143: 1–20.
Mul, D., W. Oostdijk and S.L. Drop, 2002. "Early puberty in adopted children." *Hormone Research in Pediatrics* 57, 1–9.
Parent A-S, Teilmann G, A. Juul, N.E. Skakkebaek, J. Toppari J, and J-P Bourguignon, 2003. "The timing of normal puberty and the age limits of sexual precocity: variations around the world, secular trends, and changes after migration." *Endocr Rev* 24:668–693
Parker, S. W and T. Vogl. 2018. "Do Conditional Cash Transfers Improve Economic Outcomes in the Next Generation? Evidence from Mexico." *NBER Working Paper* No. 24303.
Raven, J.C., Court, J.H. and Raven, J. 1984. *Manual for Raven's Progressive Matrices and Vocabulary Scales. Section 2: Coloured Progressive Matrices*. London: H. K. Lewis.
Stecklov, G., P. Winters, J. Todd, and F. Regalia. 2007. "Unintended effects of poverty programmes in childbearing in less developed countries: Experimental evidence from Latin America," *Population Studies* **61**(2): 125–140.
Todd, J. E., P. Winters, and G. Stecklov. 2012. "Evaluating the impact of conditional cash transfer programs on fertility: The case of the *Red de Protección Social* in Nicaragua," *Journal of Population Economics* **25**: 267–290.
World Bank. 2015. *The State of Social Safety Nets 2015*. Washington, DC: World Bank.
World Bank, 2020. *Cost-effective Approaches to Improve Global Learning*. Recommendations of the Global Education Evidence Advisory Panel, Washington, DC: World Bank.
Young, A., 2019. "Channeling Fisher: Randomization Tests and the Statistical Insignificance of Seemingly Significant Experimental Results", *Quarterly Journal of Economics* 134(2): 557–98.

TABLE 1: 2010 LONG-TERM EXPERIMENTAL DIFFERENTIALS FOR ALL FAMILIES OF OUTCOMES, 7–12 AND 9–12 COHORTS, MALES AND FEMALES

	Labor Market Participation Family Z-Score	Earnings Family Z-Score		Schooling Family Z-score	Learning Family Z-Score	Fertility and Marriage Family Z-Score	Socio-Emotional Family Z-score
		Rank	Earnings (5 % Trim)				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: 7–12-Year-old Males and Females							
ITT	0.174***	0.140***	0.120***	0.064*	0.061	-	0.007
	(0.046)	(0.045)	(0.044)	(0.034)	(0.050)		(0.023)
N	2,897	2,897	2,873	2,898	2,711		2,690
Panel B: 9–12-Year-old Males and Females							
ITT	0.198***	0.141***	0.140***	0.097***	0.092*	-	-0003
	(0.055)	(0.048)	(0.050)	(0.030)	(0.048)		(0.030)
N	1,894	1,894	1,875	1,895	1,734		1,721
Panel C: 9–12-Year-old Males							
ITT	0.272***	0.194***	0.192***	0.098**	0.183**	-0.059	0.053
	(0.075)	(0.057)	(0.067)	(0.043)	(0.070)	(0.064)	(0.039)
FWER p	[0.005]	[0.005]	[0.011]	[0.023]	[0.014]	[0.116]	[0.067]
RI p	[0.002]	[0.001]	[0.008]	[0.036]	[0.030]	[0.442]	[0.257]
N	1,006	1,006	997	1,007	907	907	900
Panel D: 9–12-Year-old Females							
ITT	0.169**	0.116*	0.104*	0.096**	-0.005	-0.167***	-0.053
	(0.074)	(0.061)	(0.060)	(0.040)	(0.057)	(0.060)	(0.050)
FWER p	[0.066]	[0.072]	[0.076]	[0.066]	[0.361]	[0.066]	[0.141]
RI p	[0.038]	[0.072]	[0.098]	[0.020]	[0.936]	[0.025]	[0.325]
N	888	888	878	888	827	809	821
Panel E: Difference between Male and Female 9–12 cohorts							
P-value	0.288	0.268	0.263	0.964	0.043	0.276	0.102

Notes: Outcome data source: 2010 long-term evaluation survey. Individual-level variables in columns 1–4 measured using the 2010 household survey instrument, in columns 5 and 7 using the 2010 individual survey instrument, and in column 6 using both. Regressions include males and/or females 7–12 or 9–12 years old at the start of the program in November 2000, weighted to account for sampling and attrition providing population estimates of the study area. The differential ITT results compare early to late treatment groups in 2010. Family z-scores are calculated by averaging the z-score for the individual components in each family (shown in tables 2–6), with individual z-scores calculated using the mean and standard deviation of the late treatment group. The fertility and marriage family has different components for males and females and therefore is not combined. Controls include three-month age group indicators and indicator variables for grades completed at baseline, stratification and region. Standard errors clustered at the locality level are in parentheses. FWER p shows in brackets p-values adjusted for multiple hypotheses testing using the sharpened familywise error rate following Anderson (2008) and based on the variables included in the table. RI p shows Fisher exact p-values obtained through randomization inference using Young's (2019) randomization-t. *** p<0.01, ** p<0.05, * p<0.10.

TABLE 2: 2010 LONG-TERM EXPERIMENTAL DIFFERENTIALS FOR LABOR MARKET PARTICIPATION FAMILY COMPONENTS, 9–12 COHORT BY SEX

	Labor Market Participation Family Z-Score	Labor Market Participation Family Components			
		Worked Off Family Farm =1 (Last 12 Months)	Migrated for Work =1 (Last 12 months)	Ever Had a Salaried Non-Agricultural Job =1	Ever worked in Urban Area =1
	(1)	(2)	(3)	(4)	(5)
Panel A: Males					
ITT	0.272***	0.062***	0.093***	0.084**	0.065*
	(0.075)	(0.022)	(0.032)	(0.036)	(0.034)
N	1,006	1,006	1,006	998	998
Mean Late Treatment		0.828	0.312	0.226	0.127
Panel B: Females					
ITT	0.169**	0.069*	0.087***	0.020	0.016
	(0.074)	(0.038)	(0.024)	(0.037)	(0.031)
N	888	888	887	883	883
Mean Late Treatment		0.463	0.074	0.312	0.234

Notes: Outcome data source: 2010 long-term evaluation household survey instrument. Regressions include males or females 9–12 years old at the start of the program in November 2000, weighted to account for sampling and attrition providing population estimates of the study area. The differential ITT results compare early to late treatment groups in 2010. The family z-score is calculated by averaging the z-score for the four individual components with the average calculated even if some components are missing. Individual component z-scores are calculated using the mean and standard deviation of the late treatment group. Controls include three-month age group indicators and indicator variables for grades completed at baseline, stratification and region. Standard errors clustered at the locality level are in parentheses. *** p<0.01, ** p<0.05, * p<0.10.

TABLE 3: 2010 LONG-TERM EXPERIMENTAL DIFFERENTIALS FOR EARNINGS FAMILY COMPONENTS, 9–12 COHORT BY SEX

	Earnings Family Z-Score	Earnings Family Components (all unconditional)			
		Earnings Per Month Worked (Last 12 Months)	Annual Earnings (Last 12 Months)	Maximum Monthly Earnings (Last 12 Months)	Maximum Monthly Non-agricultural Salary (Ever)[1]
	(1)	(2)	(3)	(4)	(5)
Panel A: Rank of Earnings, Males					
ITT	0.194***	41.780**	25.568	43.899**	49.313**
	(0.057)	(19.471)	(18.493)	(19.290)	(19.684)
N	1,006	1,006	1,006	998	1,006
Mean Late Treatment		497.2	502.5	486.9	498.3
Panel B: Earnings (C$) 5% Trim, Males					
ITT	0.192***	201.152***	595.013	211.421***	142.260*
	(0.067)	(63.624)	(619.322)	(69.318)	(71.919)
N	997	956	956	956	955
Mean Late Treatment		1436	8222	1619	227.5
Panel C: Rank of Earnings, Females					
ITT	0.116*	33.758**	28.195*	4.148	31.313*
	(0.061)	(15.706)	(14.011)	(15.038)	(15.544)
N	888	888	888	883	888
Mean Late Treatment		414.8	417.4	434.9	415.5
Panel D: Earnings (C$) 5% Trim, Females					
ITT	0.104*	97.623	95.031	134.973**	16.071
	(0.060)	(62.401)	(328.513)	(56.948)	(43.562)
N	878	848	856	848	839
Mean Late Treatment		464.2	2628	492.9	309.4

Notes: Outcome data source: 2010 long-term evaluation household survey instrument. Regressions include males or females 9–12 years old at the start of the program in November 2000, weighted to account for sampling and attrition providing population estimates of the study area. The differential ITT results compare early to late treatment groups in 2010. Earnings include wage work off the family farm. Earnings in panels B and D are trimmed at the top 5% of values. Earnings shown in Nicaraguan Cordobas (C$); the exchange rate was approximately 20 C$ per U.S. dollar in 2010. The family z-score is calculated by averaging the z-score for the four individual components with the average calculated even if some components are missing. Individual component z-scores are calculated using the mean and standard deviation of the late treatment group. Controls include three-month age group indicators and indicator variables for grades completed at baseline, stratification and region. Standard errors clustered at the locality level are in parentheses. *** p<0.01, ** p<0.05, * p<0.10.

[1] Because relatively few men ever held a non-agricultural salaried job, the unconditional mean in Panel B, column 5 is much lower than earnings per month worked in column 1.

TABLE 4: 2010 LONG-TERM EXPERIMENTAL DIFFERENTIALS FOR SCHOOLING, LEARNING AND COGNITION FAMILY COMPONENTS, 9–12 COHORT BY SEX

	Schooling Family Z-Score	Schooling Family Components			Read and Write =1	Learning Families			Mixed Cognition and Learning	Cognition (Raven)
		Grades Completed	Completed Grade 4=1	Enrolled =1		Math and Spanish	Math	Spanish		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: Males										
ITT	0.098**	0.288*	0.035	0.045**	0.052**	0.183**	0.160**	0.204**	0.113	-0.016
	(0.043)	(0.167)	(0.024)	(0.021)	(0.021)	(0.070)	(0.069)	(0.081)	(0.082)	(0.095)
N	1,007	1,006	1,006	1,005	1,007	907	905	907	906	906
Mean Late Treatment		5.498	0.747	0.181	0.870					
Panel B: Females										
ITT	0.096**	0.177	0.066***	0.022	0.001	-0.005	-0.000	-0.010	-0.047	-0.011
	(0.040)	(0.141)	(0.020)	(0.029)	(0.014)	(0.057)	(0.055)	(0.065)	(0.061)	(0.088)
N	888	888	888	885	888	827	827	826	826	826
Mean Late Treatment		6.758	0.825	0.296	0.956					

Notes: Outcome data source: 2010 long-term evaluation survey. Individual-level variables in columns 1–5 measured using the 2010 household survey instrument and in columns 6–10 using the 2010 individual survey instrument. Regressions include males or females 9–12 years old at the start of the program in November 2000, weighted to account for sampling and attrition providing population estimates of the study area. The differential ITT results compare early to late treatment groups in 2010. Family z-scores are calculated by averaging the z-score for the individual components with the average calculated even if some components are missing. Individual component z-scores for tests are calculated using the mean and standard deviation of the late treatment group. Controls include three-month age group indicators and indicator variables for grades completed at baseline, stratification and region. Standard errors clustered at the locality level are in parentheses. *** p<0.01, ** p<0.05, * p<0.10.

TABLE 5: 2010 LONG-TERM EXPERIMENTAL DIFFERENTIALS FOR FERTILITY AND MARRIAGE FAMILY COMPONENTS, 9–12 COHORT BY SEX

	Fertility and Marriage Family Z-Score	Fertility and Marriage Family Z-Score (Excluding BMI)	Fertility and Marriage Family Components					Attended CCT Workshop on Reproductive Health =1	Knows What a Pap Test is =1
			Had Sex by Age 15=1	Ever Married =1	Any Children =1	Age of Menarche	Body Mass Index (BMI)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Males									
ITT	-0.059		0.080**	-0.094**	-0.048			-0.192***	
	(0.064)		(0.035)	(0.042)	(0.038)			(0.032)	
N	907		875	907	875			825	
Mean Late Treatment			0.269	0.311	0.225			0.863	
Panel B: Females									
ITT	-0.167***	-0.166**	-0.109**	-0.039	-0.064	0.249**	-0.656***	-0.058**	-0.063*
	(0.060)	(0.063)	(0.041)	(0.043)	(0.047)	(0.119)	(0.236)	(0.028)	(0.037)
N	809	809	809	809	809	806	765	749	792
Mean Late Treatment			0.287	0.612	0.527	13.11	23.64	0.815	0.751

Notes: Outcome data source: 2010 long-term evaluation survey. Variables are measured using the 2010 individual survey instrument with the exception of marital status measured in the household instrument. Regressions include males or females 9–12 years old at the start of the program in November 2000, weighted to account for sampling and attrition providing population estimates of the study area. The differential ITT results compare early to late treatment groups in 2010. Age of menarche is reversed when it is included in the fertility and marriage family. BMI is not available for men. BMI included only for women not pregnant at the time of measurement. The family z-scores are calculated by averaging the z-score for the individual components with the average calculated even if some components are missing. Individual component z-scores are calculated using the mean and standard deviation of the late treatment group. Controls include three-month age group indicators and indicator variables for grades completed at baseline, stratification and region. Standard errors clustered at the locality level are in parentheses. *** p<0.01, ** p<0.05, * p<0.10.

TABLE 6: 2010 LONG-TERM EXPERIMENTAL DIFFERENTIALS FOR SOCIO-EMOTIONAL FAMILY COMPONENTS, 9–12 COHORT BY SEX

	Family Z-Score	Socio-emotional Family Components			
		Optimism	Positive Self-perception	Negative Self-perception	Stress
	(1)	(2)	(3)	(4)	(5)
Panel A: Males					
ITT	0.053	0.287***	0.249**	0.155*	0.170**
	(0.039)	(0.078)	(0.093)	(0.086)	(0.071)
N	900	900	900	900	900
Panel B: Females					
ITT	-0.053	0.022	-0.012	0.153*	0.069
	(0.050)	(0.064)	(0.092)	(0.084)	(0.083)
N	821	821	821	821	821

Notes: Outcome data source: 2010 long-term evaluation individual survey instrument. Regressions include males or females 9–12 years old at the start of the program in November 2000, weighted to account for sampling and attrition providing population estimates of the study area. The differential ITT results compare early to late treatment groups in 2010. Socio-emotional components are the first four factors resulting from exploratory factor analysis of all socio-emotional questions. The family z-score is calculated by averaging the z-score for the four individual components (after reversing signs for negatively oriented self-perception and stress). Individual component z-scores are calculated using the mean and standard deviation of the late treatment group. Controls include three-month age group indicators and indicator variables for grades completed at baseline, stratification and region. Standard errors clustered at the locality level are in parentheses. *** p<0.01, ** p<0.05, * p<0.10.

TABLE 7: 2010 LONG-TERM EXPERIMENTAL DIFFERENTIALS FOR ALL FAMILIES, TWO-YEAR COHORTS BY SEX

	Labor Market Participation Family Z-Score	Earnings Family Z-Score: Rank	Earnings Family Z-Score: Absolute (5 % Trim)	Schooling Family Z-Score	Learning Family Z-Score	Fertility and Marriage Family Z-score	Socio-Emotional Family Z-Score
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: 7–8-Year-old Males							
ITT	0.131	0.160	0.097	-0.008	-0.041	-0.117	-0.013
	(0.091)	(0.096)	(0.091)	(0.079)	(0.080)	(0.153)	(0.045)
N	498	498	496	498	499	548	492
Panel B: 9–10-Year-old Males							
ITT	0.213***	0.173*	0.183**	0.126**	0.173*	-0.030	0.001
	(0.079)	(0.090)	(0.079)	(0.058)	(0.092)	(0.072)	(0.061)
N	466	466	463	467	430	430	427
Panel C: 11–12-Year-old Males							
ITT	0.341***	0.220***	0.208*	0.066	0.187**	-0.062	0.094**
	(0.110)	(0.071)	(0.106)	(0.078)	(0.079)	(0.082)	(0.045)
N	540	540	534	540	477	477	473
Panel D: 7–8-Year-old Females							
ITT	0.097	0.076	0.060	0.046	0.079	-0.001	0.079
	(0.101)	(0.097)	(0.102)	(0.073)	(0.096)	(0.079)	(0.062)
N	505	505	502	505	478	473	477
Panel E: 9–10-Year-old Females							
ITT	0.119	0.120	0.146*	0.100	0.085	-0.240***	0.011
	(0.077)	(0.079)	(0.081)	(0.077)	(0.069)	(0.076)	(0.061)
N	403	403	399	403	380	374	377
Panel F: 11–12-Year-old Females							
ITT	0.184*	0.088	0.026	0.100*	-0.086	-0.095	-0.118*
	(0.099)	(0.087)	(0.088)	(0.058)	(0.069)	(0.069)	(0.060)
N	485	485	479	485	447	435	444
Panel G: P-value of Difference between 7–8 and 9–12 cohorts							
Males	0.002	0.019	0.049	0.012	0.012	0.329	0.390
Females	0.068	0.193	0.272	0.024	0.911	0.032	0.939

Notes: Outcome data source: 2010 long-term evaluation survey. Individual-level variables in columns 1–4 measured using the 2010 household survey instrument, in columns 5 and 7 using the 2010 individual survey instrument, and in column 6 using both. Regressions include males or females of the indicated ages at the start of the program in November 2000, weighted to account for sampling and attrition providing population estimates of the study area. The differential ITT results compare early to late treatment groups in 2010. Family z-scores are calculated by averaging the z-score for the individual components in each family, with individual z-scores calculated using the mean and standard deviation of the late treatment group. The fertility family is different for males versus females. Controls include three-month age group indicators and indicator variables for grades completed at baseline, stratification and region. Standard errors clustered at the locality level are in parentheses. *** p<0.01, ** p<0.05, * p<0.10.

TABLE 8: 2002 AND 2004 AVERAGE ENROLLMENT RATES, TWO-YEAR COHORTS BY SEX

	Males			Females		
	7–8	9–10	11–12	7–8	9–10	11–12
2002 Early Treatment	1.00	0.97	0.92	0.95	1.00	0.96
2004 Late Treatment	0.95	0.89	0.58	0.97	0.86	0.68

Notes: Outcome data source: 2002 and 2004 short-term evaluation surveys. Means are based on males or females with the indicated age at the start of the program in November 2000, weighted to account for sampling in the short-term evaluation survey providing population estimates of the study area.

TABLE 9: 2002 AND 2004 SHORT-TERM EXPERIMENTAL DIFFERENTIALS FOR EDUCATION, 9–12 COHORT BY SEX

	Grades Completed (years)	Completed Grade 4 =1	Enrolled =1	Regularly attended School > 85% of Time =1	Read and Write =1
	(1)	(2)	(3)	(4)	(5)
Panel A: Males 2002					
ITT	0.361***	0.053*	0.182***	0.360***	0.150***
	(0.094)	(0.031)	(0.042)	(0.055)	(0.034)
N	475	475	475	475	475
Mean LT	2.396	0.277	0.733	0.544	0.735
Panel B: Males 2004					
ITT	0.487***	0.086*	-0.049	-0.100	0.124***
	(0.155)	(0.045)	(0.063)	(0.066)	(0.029)
N	458	458	458	458	458
Mean LT	3.585	0.536	0.626	0.564	0.815
Panel C: Females 2002					
ITT	0.266***	0.048*	0.099***	0.177***	0.096***
	(0.048)	(0.025)	(0.019)	(0.027)	(0.029)
N	450	450	450	450	450
Mean LT	2.952	0.343	0.875	0.781	0.845
Panel D: Females 2004					
ITT	0.573***	0.184***	-0.141***	-0.149**	0.032
	(0.117)	(0.036)	(0.051)	(0.060)	(0.022)
N	394	394	394	394	394
Mean LT	4.357	0.665	0.766	0.682	0.930

Notes: Outcome data source: 2002 and 2004 short-term evaluation surveys. Regressions include males or females ages 9–12 at the start of the program in November 2000, weighted to account for sampling in the short-term evaluation survey providing population estimates of the study area. (Results for other age groups shown in Appendix Table A7). The differential ITT results compare early to late treatment groups. The late treatment group began receiving the program in 2003 so that the ITT captures absolute effects in 2002, and short-term differential effects in 2004. Mean LT refers to mean in the late treatment group. Grades completed reflect the completed school year prior to each survey year. Attended school for more than 85% of the time is measured over the previous month and is zero for those who were not enrolled in school at the time. Controls include three-month age group indicators and indicator variables for grades completed at baseline, stratification and region. Standard errors clustered at the locality level are in parentheses. *** p<0.01, ** p<0.05, * p<0.10.

TABLE 10: 2002 AND 2004 SHORT-TERM EXPERIMENTAL DIFFERENTIALS FOR CONSUMPTION, HOUSEHOLDS OF 9–12 COHORT CHILDREN BY SEX

	Log Food Consumption per Capita	Share of Food Consumption On:			
		Animal Protein	Fruit and Vegetables	Staples	Other Food
	(1)	(2)	(3)	(4)	(5)
Panel A: Males 2002					
ITT	0.354***	0.046**	0.041***	-0.082***	-0.005
	(0.085)	(0.018)	(0.007)	(0.020)	(0.012)
N	500	500	500	500	500
Mean LT	7.309	0.206	0.054	0.534	0.206
Panel B: Males 2004					
ITT	-0.273***	-0.001	-0.000	0.009	-0.005
	(0.068)	(0.016)	(0.008)	(0.016)	(0.012)
N	490	494	494	494	494
Mean LT	8.012	0.201	0.085	0.512	0.196
Panel C: Females 2002					
ITT	0.345***	0.054***	0.042***	-0.074***	-0.023**
	(0.083)	(0.013)	(0.007)	(0.018)	(0.009)
N	475	475	475	475	475
Mean LT	7.343	0.216	0.059	0.508	0.216
Panel D: Females 2004					
ITT	-0.133**	-0.037**	0.004	0.023	0.015*
	(0.065)	(0.015)	(0.011)	(0.018)	(0.008)
N	459	465	465	465	465
Mean LT	7.827	0.179	0.089	0.530	0.201

Notes: Outcome data source: 2002 and 2004 short-term evaluation surveys. Regressions include households with females ages 9–12 at the start of the program in November 2000, weighted to account for sampling in the short-term evaluation survey providing population estimates of households with boys (girls) this age in the study area. The differential ITT results compare early to late treatment groups. The late treatment group began receiving the program in 2003 so that the ITT captures absolute effects in 2002, and short-term differential effects in 2004. Mean LT refers to mean in the late treatment group. Food consumption is measured as the total value of purchased and otherwise obtained food using a comprehensive food consumption module (Maluccio and Flores, 2005). Controls include three-month age group indicators and indicator variables for grades completed at baseline, stratification and region. Standard errors clustered at the locality level are in parentheses. *** p<0.01, ** p<0.05, * p<0.10.