Tania Barham
Karen Macours
John A Maluccio
Ferdinando Regalia
Veronica Aguilera
Miriam Enoe Moncada

# Assessing long-term impacts of conditional cash transfers on children and young adults in rural Nicaragua

July 2014

Impact Evaluation Report 17

Social protection

3ie International Initiative for Impact Evaluation

**About 3ie**

The International Initiative for Impact Evaluation (3ie) was set up in 2008 to meet growing demand for more and better evidence of what development interventions in low- and middle-income countries work and why. By funding rigorous impact evaluations and systematic reviews and by making evidence accessible and useful to policymakers and practitioners, 3ie is helping to improve the lives of people living in poverty.

**3ie Impact Evaluations**

3ie-supported impact evaluations assess the difference a development intervention has made to social and economic outcomes. 3ie is committed to funding rigorous evaluations that include a theory-based design, use the most appropriate mix of methods to capture outcomes and are useful in complex development contexts.

**About this report**

3ie accepted the final version of this report, *Assessing long-term impacts of conditional cash transfers on children and young adults in rural Nicaragua*, as partial fulfilment of requirements under grant OW2.216 issued under Open Window 2. 3ie has copyedited and formatted the content for publication. Due to unavoidable constraints at the time of publication, a few of the tables or figures may be less than optimal. All of the content is the sole responsibility of the authors and does not represent the opinions of 3ie, its donors or its Board of Commissioners. Any errors and omissions are also the sole responsibility of the authors. All affiliations of the authors listed in the title page are those that were in effect at the time the report was accepted. Any comments or queries should be directed to the corresponding author, Karen Macours at Karen.macours@psemail.eu

Funding for this impact evaluation was provided by 3ie's donors, which include UKaid, the Bill & Melinda Gates Foundation, Hewlett Foundation and 12 other 3ie members that provide institutional support. A complete listing is provided on the 3ie website at www.3ieimpact.org/about-us/3ie-members.

Suggested citation: Barham, T, Macours, K, Maluccio, JA, Regalia, F, Aguilera, V and Moncada, ME, 2014. *Assessing long-term impacts of conditional cash transfers on children and young adults in rural Nicaragua. 3ie Impact Evaluation Report 17.* New Delhi: International Initiative for Impact Evaluation (3ie).

3ie Impact Evaluation Report Series executive editors: Jyotsna Puri and Beryl Leach
Managing editors: Stuti Tripathi, Thomas de Hoop and Shagun Sabarwal
Assistant managing editor: Kanika Jha
Production managers: Lorna Fray and Omita Goyal
Assistant production manager: Rajesh Sharma
Copy editor: Lucy Southwood
Proofreader: Proteeti Banerjee
Cover design: John F McGill
Cover photo: Arne Hoel/ World Bank
Printer: Via Interactive

© International Initiative for Impact Evaluation (3ie), 2014

# Assessing long-term impacts of conditional cash transfers on children and young adults in rural Nicaragua

Tania Barham
University of Colorado, Boulder

Karen Macours
Paris School of Economics

John A Maluccio
Middlebury College

Ferdinando Regalia
Inter-American Development Bank

Veronica Aguilera
Centro de Investigaciones y Estudios Rurales y Urbanos de Nicaragua

Miriam Enoe Moncada
Centro de Investigaciones y Estudios Rurales y Urbanos de Nicaragua

**3ie Impact Evaluation Report 17**
**July 2014**

International Initiative for Impact Evaluation

## Acknowledgements

Research discussed in this report has been funded by the International Initiative for Impact Evaluation, Inc. (3ie) through the Global Development Network (GDN) (OW2.216), the Inter-American Development Bank (IDB) and the National Science Foundation (SES 11239945 and 1123993). The views expressed in this report are not necessarily those of 3ie or its members. We are indebted to the entire survey team from Centro de Investigaciones y Estudios Rurales y Urbanos de Nicaragua (CIERUNIC) for their excellent data collection and dogged persistence in tracking. We also gratefully acknowledge members of the RPS programme team (in particular, Leslie Castro, Carold Herrera and Mireille Vijil) for discussions regarding the evaluation, and Emma Sanchez Monin for facilitating the data collection process. We thank Teresa Molina, Olga Larios and Gisella Kagy for help with data preparation and Vincenzo di Maro for numerous contributions to this research project.

## Abstract

Conditional cash transfer (CCT) programmes, which offer families cash grants conditional on attendance at school or preventive health visits, have expanded rapidly over the past decade. Although there is substantial evidence of their positive short-term impacts on education, health and poverty reduction, little is known about their long-term effects on the human capital and labour market outcomes they target to break the inter-generational transmission of poverty.

The research builds on a previously implemented randomised CCT programme in Nicaragua that operated from 2000 to 2005, the *Red de Protección Social* [social protection network] (RPS). RPS targeted rural households in impoverished regions of Nicaragua and had two household-level transfers – one for health and nutrition and the other for schooling – in households with eligible school-age children. Each of two experimental groups received monetary benefits for only three years, and the programme was randomly phased-in with the early treatment group starting in 2000 and the late treatment group in 2003. Our principal identification strategy exploits the randomised design and difference in timing of benefits in the two groups of beneficiaries, as well as age-specific conditionalities of various components of the programme. Since both experimental groups eventually received programme benefits, we examine the differential effect of the programme's timing. For some analysis and outcomes, where it is suitable, we also use non-experimental methods to determine the absolute effect.

We investigate whether a well-implemented programme that led to large improvements in health, nutrition and education in the short run also resulted in improved cognitive functioning, schooling, learning and labour market outcomes 10 years after the programme start. We also explore the programme's sustainability via analysis of household-level expenditures. Given the research questions and the potential (but unknown) relationship between cognitive development, education and migration, it was paramount to our study to minimise attrition. As such, extensive migrant tracking was a key component of the research design.

We find that boys exposed *in utero* and during the first two years of life, have better cognitive outcomes when they are 10 years old than those exposed in their second year of life or later. These results are important from a policy perspective as they confirm that interventions that improve nutrition and/or health during the first 1,000 days of life can have lasting positive impacts on cognitive development for children. However, there are no differential impacts on anthropometrics, despite short-term differences resulting from the programme, demonstrating that complete catch-up in anthropometrics was possible. We also find catch-up in physical growth for the same age girls.

For schooling and learning outcomes, we focus on school-age cohorts at risk of dropping out when the intervention began. For boys aged 9–12 in 2000 (and thus 19–22 in 2010), the short-term programme effect of a half-grade increase in schooling was sustained into early adulthood, seven years after the end of the programme. In addition, there were significant and substantial gains in both maths and language achievement scores, an approximately one quarter standard deviation increase in learning outcomes for the now young men. Hence in Nicaragua, schooling and achievement gains coincided, implying important longer-term returns to CCT programmes. Recently, questions have been raised about the potential of schooling programmes in general, and of CCTs in particular,

to affect learning outcomes. These results provide a clear positive answer to this question, an important conclusion for policymakers.

Moreover, these gains in achievement yielded dividends in the labour market. We find that those boys of the same cohort in the early treatment group have higher earnings in the labour market than those in the late treatment group. They are more likely to work outside the home, and indeed the earning gains are obtained through more seasonal migration. Hence even in a context where work in good non-agriculture jobs is difficult to find, the programme led to improvements in income.

Our assessment of the impact of RPS on schooling, learning and labour market outcomes hinges importantly on the timing of the programme transfers and conditions relative to when children were at high risk of dropping out. These patterns were different for girls, who tended to drop out of school two to three years later than boys. While we have less detailed information on this older cohort, we find that girls aged 13–14 at the start of the programme had nearly three quarters of a grade more schooling in early versus late treatment.

Schooling, learning and labour market outcomes for young adults are clear indications of the programme's persistent or sustainable effects. Our findings on other indicators of sustainability, including changes in expenditure patterns, are less clear.

Overall, the results clearly show that CCTs can lead to sustainable gains in human capital accumulation, and that these remain long after households exit the programme. This suggests they have potential to lead to long-term poverty reduction.

## Contents

## List of figures and tables



## Appendix figures and tables

## Abbreviations and acronyms

| | |
|---|---|
| CCT | conditional cash transfer |
| CIERUNIC | Centro de Investigaciones y Estudios Rurales y Urbanos de Nicaragua |
| GDN | Global Development Network |
| GDP | gross domestic product |
| ICC | intra-cluster correlations |
| IDB | Inter-American Development Bank |
| INCAP | Guatemalan Institute for Nutrition in Central America and Panama |
| ITT | intention to treat |
| LLR | local linear regression |
| RCT | randomised controlled trial |
| RPS | *Red de Protección Social* (social protection network) |
| sd | standard deviation |

# 1. Introduction

Conditional cash transfer (CCT) programmes have expanded rapidly over the past decade, and have been implemented in more than 30 countries worldwide (Fizsbein and Schady 2009; IEG 2011). In many Latin American countries, CCTs have become the largest national anti-poverty programme; they also have been expanding in Africa, South Asia, East Asia and even in the US. Numerous evaluations, many based on rigorous experimental designs, leave little doubt that such programmes can reduce poverty and increase schooling, nutrition and health outcomes – in the short term. But evidence is notably lacking on whether these short-term gains eventually translate into longer-term human capital benefits that might ultimately be needed to fully justify the continuation of such programmes. In this report, we present evidence on the impact of a CCT programme 10 years after it started and, in doing so, provide unique evidence on the potential for long-term gains.

An examination of the long-term benefits of CCTs is directly linked to their specific objectives, since improving children's health and nutrition is often argued to be important not only for immediate well-being, but also because they may reduce long-term poverty through improvements in cognitive development, health, educational achievements and ultimately, in labour market opportunities (Strauss and Thomas 2008; Heckman 2006). Economic theories of human capital development rely on this postulated link and are part of the rationale for social programmes in both the developed and the developing worlds.

Yet, there is limited evidence of the long-term effects of early childhood health and nutrition interventions on later cognitive functioning, children's learning in school and labour market outcomes (Hoddinott *et al.* 2008; Maluccio *et al.* 2009; Barham 2012).[1] Nor is it well understood how important the exact timing of health and nutrition interventions are for effectiveness. This is because it is unclear whether the benefits from interventions occurring in the earliest years of life – for example, before the age of two, continue (Pollitt *et al.* 1993) or fade-out (Garces *et al.* 2002) – or whether interventions after the age of two can help reverse the effects of prior deficits (Walker *et al.* 2007). The issue of fade-out is particularly pertinent in developing countries, where there are many competing health risks and a greater frequency of shocks, coupled with limited capacity to smooth consumption. It is also important since many of the countries implementing these programmes are of low or medium income and therefore may not be able to afford to provide the programmes on a continuing basis or to children of all ages.

Like childhood health and nutrition, education is also regarded as an important path to future poverty reduction. While there is substantial literature assessing, for example, the effect of schooling on wages and age of first birth, there is concern that many previous studies suffer from endogeneity bias due to the difficulty of controlling for unobservables such as ability. As a result, the true causal effects are unknown (Griliches 1977; Card 1999, 2001; Oliver 1999; Thomas 1999; Glewwe 2002). In developing countries, further

[1] For *Oportunidades* in Mexico, Gertler *et al.* (2012) show longer-term impacts on household investment in economic activities. For a different CCT programme in Nicaragua, Macours, Schady, and Vakis (2012) find effects on early childhood cognitive outcomes and sustained changes in human capital investment behaviour two years after the end of the programme, while Macours, Premand, and Vakis (2012) show that complementary productive interventions added to that CCT helped households protect themselves against shocks, even after the programme ended.

substantial concerns about low-quality schooling and labour market imperfections leave policymakers uncertain as to whether programmes that increase schooling will actually result in improved learning and hence improved labour market outcomes.

Our project builds on the *Red de Protección Social* [Social protection network] (RPS) – a previously implemented randomised CCT programme that operated in Nicaragua from 2000 to 2005 – to address several shortcomings in the literature. We investigate whether a well-implemented programme that led to large short-term improvements in health, nutrition and education also resulted in improved cognitive functioning, learning and labour market outcomes 10 years on, for those who received the programme at critical ages. In this report, we refer to these as long-term effects since they are in contrast to a more typical one-to-two-year short-term evaluation, and because some outcomes examined represent permanent changes (such as higher grades attained).

Our principal strategy for answering these questions exploits the randomised design and difference in timing of benefits in two groups of beneficiaries, as well as age-specific conditionalities of various programme components. Since both experimental groups eventually receive programme benefits, we examine the programme's differential effect. For some analysis and outcomes, particularly where there is sufficient sample size, we also use non-experimental methods to determine the programme's absolute effect (see Appendix G). We examine a comprehensive set of cognitive functioning and educational achievement (or learning) outcomes. We also investigate whether the programme affected reproductive behaviour, drawing on qualitative work that was an integral part of the research project. The qualitative work also sheds light on the sustainability of the shifts in households' human capital investment and consumption behaviour. We further investigate this using quantitative household expenditure data.

Given the research questions and the potential (but unknown) relationship between cognitive development, education, health and migration, it was paramount to our study to minimise attrition. Consequently, migrants were tracked throughout the country, and to neighbouring Costa Rica, the dominant destination for international migrant labour from Nicaragua. Extensive tracking was a key component of the research design.[2]

In Section 2, we provide further context for the study. In Section 3, we describe the intervention and its relation to other CCTs, as well as the programme implementation (which was not a direct aspect of this research project, having ended five years earlier). Section 4 describes the tracking of migrants in our study, while Section 5 summarises our principal findings to date. We conclude in Section 6 with a discussion of policy implications.

## 2. Context

Prior to discussing the long-term results, in this section we review the short-term evidence of CCTs in general, and of the Nicaraguan CCT in particular. The short-term impacts provide key contextual factors for interpreting the longer-term benefits. We also discuss related relevant literature.

[2] Such tracking was the primary objective and use of the funding provided by 3ie; the main data collection was funded by the IDB.

## 2.1 CCTs and the CCT programme in Nicaragua

CCT programmes have led to important short-term (usually within one or two years) benefits in a number of areas including education, health and nutrition. School enrolment rates have increased in many different countries and regions as a result of these programmes. For example, in Ecuador, a CCT-type programme targeting children aged 6–17 increased enrolment for this age group by approximately 10 per cent (Schady and Araujo 2008). In Mexico, enrolment increased by 5 per cent for primary school students (it is not higher largely because primary-school enrolment in Mexico was already high) and by approximately 12 per cent for grades 6–8 (Schultz 2004). In terms of health and nutrition, CCTs have substantially increased healthcare utilisation rates (Bautista-Arredondo *et al.* 2006), increased immunisation against major childhood diseases (Barham and Maluccio 2009), reduced episodes of acute diarrhoea (Attanasio *et al.* 2005), improved the nutrition status of children (Gertler 2004; Maluccio and Flores 2005; Rivera *et al.* 2004; Attanasio *et al.* 2005; Behrman and Hoddinott 2005) and reduced infant mortality (Barham 2011).

The short-term effects, after two years, were particularly noticeable for RPS in Nicaragua. Enrolment of 7–13-year-olds rose by approximately 13 per cent from 2000 to 2002 and the effects on regular school attendance were even greater (Maluccio and Flores 2005; Maluccio *et al.* 2010). The programme led to a 5 per cent fall in stunted children under five and there were substantial increases in food expenditure (Maluccio and Flores 2005). The percentage of children aged 0–3 weighed in the past six months increased by 13 per cent, and the number of two-year-olds fully vaccinated against all the major childhood diseases increased by 15 per cent (Barham and Maluccio 2009). Consequently, there were substantial short-term effects on young children, particularly those under three and of primary school age. If such improvements in health, nutrition and education lead to improved cognitive development, health, educational achievements and labour market opportunities later in life, as is suggested by the theories of human capital development, RPS should be effective at reducing long-term poverty for these individuals.

To date, there is little evidence on the long-term effects of CCTs, partly because the first large-scale programme only began in 1997. Much of the evidence that does exist is from the Mexican programme *Progresa/Oportunidades*. *Oportunidades* is an on-going programme so long-term analyses exploit an 18-month difference in length of exposure between original treatment and control groups, as the control group became eligible for the programme a year and a half after the treatment group. After 5.5 years, the number of grades attained increased for rural 15–21-year-olds, but differential exposure to the programme had no effect on achievement tests and had insignificant or negative (consistent with ongoing schooling) effects on the probability of working (Behrman *et al.* 2009, 2011). In addition, while longer exposure to the programme during early childhood is related to improved behaviour and greater growth among children of non-educated mothers, it did not yield significant differences in cognitive or language skills (Fernald *et al.* 2009).

The lack of significant positive results for *Oportunidades* is surprising in light of its substantial short-term effects. Two aspects of these follow-up analyses are likely to have played a role in the findings, both of which we address in our study. First, the differential exposure period – 18 months – may have been too short to generate large differences in measured outcomes. Second, there was substantial attrition between 1997 and 2003,

mainly due to migration—approximately 40 per cent for the 15–21-year-old age group. Moreover, a number of pre-programme characteristics are correlated with this migration (Parker *et al.* 2008).

### 2.2 Broader evidence on the effect of health, nutrition and education interventions on later life outcomes

The findings from our study are of interest beyond the literature on CCTs. In the US and other developed countries, there is growing interest in investing resources in disadvantaged children at an early stage in life. For example, Heckman (2006) draws on a wide body of evidence from economics, psychology and neuroscience to argue that returns to such investments are much higher than those made later in life. However, the empirical base for these arguments is not as deep as is often presumed and evidence from developing countries that good health and nutrition in children under five matters for later life outcomes such as cognitive functioning and income is scarce and mixed (Walker *et al.* 2007). For example, a study in Jamaica that randomised 9–24-month-olds to receive a nutritional supplement found a significant positive impact of the supplement on child development two years after the programme, but no statistically significant effects when the children were aged 7–8 (Grantham-McGregor *et al.* 1997). In contrast, results from another nutritional supplement programme run by the Guatemalan Institute for Nutrition in Central America and Panama (INCAP) which randomised four villages into two groups – one which received a high-energy protein drink and the other a placebo sugar drink – found that children who were exposed to the intervention before the age of three experienced a quarter of a standard deviation increase in both reading comprehension and non-verbal cognitive tests as young adults (Maluccio *et al.* 2009).

In addition, while some evidence on nutritional outcomes suggests it may be critical to intervene in the first two years of life, it is far less clear how sensitive cognitive outcomes are to improvements in health and nutrition at different points in early childhood. Indeed, evidence from another CCT in Nicaragua shows no significant effects on cognitive development for those 0–3, but positive impacts for children over three (Macours *et al*. 2012). Existing evidence is also inconclusive as to whether interventions that help improve the health and nutrition of a child early on – for example, if they benefit before or after the age of two or three – will differentially affect outcomes later in life, perhaps due to the possibility of reversing the prior deficits (Walker *et al.* 2007). This potential reversal may underlie the mixed results observed regarding the long-term effects of nutrition and health interventions. The potential for catch-up is an important theme in current research on developed countries (Almond and Currie 2010). Evidence on this topic is crucial for policymakers to determine the effectiveness of interventions targeted to different points in the child's life. The randomised design of our study and the timing of the interventions allow our study to address this gap.

Similarly, there is some evidence that education interventions need to take place before children drop out of school, as drop outs may be hard to reverse (de Janvry *et al.* 2006). This may be due in part to the hypothesised relationship between school drop out and child labour, as the opportunity cost of schooling increases with each additional year of age (Edmonds and Schady 2012). If such a relationship holds, we would expect the transfer programme to lead to increased schooling and achievement. Increased schooling levels and achievement can improve maths, language and possibly non-cognitive skills, which in turn may result in better labour market outcomes (Card 1999;

Duflo 2001; Cunha *et al.* 2006). In developing countries, however, where the quality of schools and teaching is often low (Glewwe 2002; Glewwe and Kremer 2006), the evidence is limited and mixed as to whether increases in schooling actually result in improvements in learning (Glewwe 2002). It is not even clear how well interventions that directly try to address the shortcomings in the education systems themselves improve learning (Glewwe and Kremer 2006; Glewwe *et al.* 2008).

In addition, while there is growing evidence around the short-term effects of varied education interventions on school attainment, drop outs and achievement in developing countries, less is known about the long-term effects of these interventions or on their potential labour market or other effects, such as marriage and health (Glewwe and Kremer 2006; Glewwe *et al.* 2008; Fiszbein and Schady 2009). For example, one recent study in Cambodia finds little effect on anything but grade attainment two years after a three-year secondary school scholarship programme (Filmer and Schady 2013).

# 3. Description of intervention and theory of change

## 3.1 The RPS programme[3]

The objective of RPS was to reduce both current and future poverty by providing cash transfers to poor households, conditional on health and education behaviours. The programme was geographically targeted to rural communities in six municipalities in central and northern Nicaragua, where the average pre-programme poverty rate, based on the 1998 Nicaraguan poverty line of US$1.12 per person per day, was 80 per cent. The municipalities are located in the Central region, then the poorest of the country's seven regions, and the only one that experienced worsening poverty between 1998 and 2001. On average, the transfers were approximately 18 per cent of pre-programme expenditures. The transfers were delivered every other month, paid to the female caregiver in the beneficiary household, referred to as the *titular.* They came with a strong social marketing message that the money was intended to improve nutrition and health and be invested in children's education. Separate amounts were transferred for, and different conditions applied to, the programme's health and education components.

The health conditionality included bringing children to scheduled preventive healthcare appointments and the *titular* attending bimonthly health information workshops. Health services at the scheduled visits included growth monitoring, vaccinations, iron supplements and anti-parasite medicine. The supply of healthcare was increased to ensure the programme could meet the increased demands for healthcare without reducing quality. In particular, RPS contracted and trained private health providers (Regalia and Castro 2007), who beneficiaries were required to use for fulfilment of the conditions. All services were free of charge. Providers visited programme areas on scheduled dates and delivered services in existing health facilities, community centres or private homes.

Receipt of the separate education transfer was contingent on all children in the household aged 7–13 who had not yet completed fourth grade enrolling in school and attending at least 85 per cent of the time. While the education component relied on the existing education infrastructure (primary schools) in the communities, a small supply-

[3] Sections 3.1 and 3.2 draw from Maluccio and Flores (2005).

side transfer was included for each child to provide schools with limited funds for school materials and to incentivise teachers. In exchange, teachers were required to report enrolment and attendance using specific RPS designed forms so the conditionalities could be verified. So, while the programme's education arm had a supply-side component, supply-side improvements were modest – and there were arguably constraints in educational quality. For example, many classrooms were multi-grade classrooms with limited educational resources, and chairs.

The programme was implemented by the government of Nicaragua with technical assistance and financial support from the IDB and benefited more than 30,000 families.

### 3.2 Programme targeting, take-up and the short-term evaluation design of RPS

The original randomised treatment and control evaluation design targeted the RPS interventions in 42 localities in six rural municipalities. The selected municipalities were relatively worse off in terms of health and education indicators. The 42 localities were selected out of a possible 59 using a poverty index based on information from the 1995 National Population and Housing Census. Finally, household targeting based on a census of the eligible localities in May 2000 led to 4 per cent of households being deemed ineligible for the programme (Maluccio 2009).

To rigorously evaluate the programme, the 42 targeted localities were randomised into equally sized treatment and control groups at a public lottery. The randomisation was stratified by poverty to ensure a similar poverty index between the two experimental areas. The 21 treatment localities became eligible for the programme after the census in mid-2000 and they were eligible to receive three years' worth of cash transfers. As with many randomised experiments, for ethical reasons the control group eventually received the programme. Therefore, after the treatment localities had received three years of transfers, the 21 control localities were phased in, starting in mid-2003. They also received three years' worth of cash transfers.

A census and a baseline household survey were carried out before the start of RPS in 2000 and three initial panel follow-up surveys took place in 2001, 2002 and 2004. Short-term evaluations using these data show significant and substantial improvements in children's health, nutrition and educational outcomes (Section 2.1), particularly those from more disadvantaged backgrounds.

Overall, compliance with the experimental design was high. Previous analysis of the programme has shown that the sample was balanced at baseline and that there was very little contamination of the late treatment group – in other words, households originally classified as living in control areas who received the programme starting in 2000. At the household level, programme take-up between early and late treatment localities was approximately 85 per cent. Moreover, compliance with conditionality was high and conditions were enforced, with transfers withheld when they were not fulfilled.

There was little evidence of contamination due to other programmes entering the area, or substitution bias (Heckman and Smith 1995). Because the government sanctioned the RPS evaluation, no other related large government programmes were started in the region over this period. RPS monitored potential contamination by other non-government development programmes in the household surveys and found that such

programmes were rare, with the exception of school feeding programmes, which increased from 2000 onward. Notably, there were no substantial differences in the presence of any programme by treatment group (Maluccio and Flores 2005).

In 2002, prior to the control group becoming eligible, RPS selected 21 additional localities in six neighbouring municipalities with similar levels of poverty to serve as a non-experimental comparison group. This was to enhance the evaluation of the long-term effects of the programme after the original control group was enrolled. A random sample of households from these localities was interviewed in the 2002 and 2004 follow-up surveys and also in our 2010 follow-up survey. In this report, we refer to the original randomised treatment group as the 'early treatment group', the original control group, who received the programme starting in 2003, as the 'late treatment group', and the additional localities added in 2002 that never received the programme as the 'comparison group'.

We emphasise that because both the early and late treatment groups received the programme, analysis comparing these two groups provides a differential programme effect. In contrast, comparing the late treatment group to the comparison group provides an absolute programme effect, though non-experimental methods need to be used and the results are more susceptible to bias. We describe the non-experimental methods in Appendix G and in Barham *et al.* 2012. Non-experimental methods cannot be used for the majority of the analyses, however, because the samples sizes are small.

### 3.3 Theory of change underlying RPS's long-term impact evaluation design

CCT programmes are social programmes designed to break the inter-generational continuity of poverty. Their basic premise is that a major cause of the transmission of poverty is the inability of poor households to invest in their children's human capital. To provide additional incentives to invest in human capital-enhancing activities and to alleviate possible credit constraints inhibiting such investments, cash transfers are given to families conditional upon family members engaging in behaviours to improve their health, education and nutrition status. This effectively transforms cash transfers into human capital subsidies. While CCTs are considered demand-side programmes, most CCT programmes that aim to improve the health and education status of poor populations also increase the supply of health and education services. Without an increase in supply, such programmes would likely lead to a decline in quality of services due to crowding from the increased demand.

Since there are excellent existing sources with detailed theories of change for CCT programmes (Gaarder *et al.* 2010; Baird *et al.* 2012), we focus in this section on the theory of change underlying RPS's long-term evaluation design. We start from the basic premise that: (a) all the households belonging to the two experimental groups that we use in most of the analysis at some point received the CCT programme; but (b) children in these households may or may not have benefitted from this intervention during critical stages in their childhood and youth. Hence our theory of change does not focus on the parts of the CCT programmes that may have operated on both types of households similarly – for example, the temporary lifting of liquidity constraints that might have triggered investments in physical assets – but rather on those components that might have differential effects depending on the ages of the children in the household.

In the first part of our analysis, we focus on the impact of the health and nutrition components on early childhood cognitive development and physical growth. The RPS intervention was designed to affect several early childhood risk factors in a number of ways. The food security transfer, by lifting liquidity constraints combined with strong social marketing, aimed at increasing the nutritional intake and quality of the diet of beneficiary households, and in particular their young children. Paying the transfers to mothers was further expected to increase investment in nutrition. The health supply component and health conditionalities increased access to preventive healthcare, including vaccinations, growth and development monitoring. The health information sessions increased overall health awareness and knowledge. While this package of interventions might have benefitted children of a wide age range, insights from the medical and nutritional literature indicate that exposure to healthy and nutritious environments might be particularly important for cognitive development and physical growth when children benefit *in utero* and very early childhood (often referred to as the 1,000-day window). Additional medical evidence also suggests that boys might be particularly vulnerable to the nutritional environment of the womb. Hence our theory of change hypothesises that exposure to CCTs during the first 1,000 days of a boy's life leads to irreversible advantages in early brain development, and these advantages are larger than when boys are exposed to a similar programme later in their childhood. This then in turn leads to sustained differential impacts on cognitive development, even when measured much later – for example, when the boys are 10 years old.

In the second part of the analysis, we focus on the CCT's education component. The programme provided households with financial incentives for children to enrol in and attend school, again combined with heavy social marketing on the importance of schooling. It was hoped that increased enrolment and attendance would lead to increased learning, which in turn could translate into higher labour market returns.[4] The age of exposure is again a critical element in the theory of change, both because of age-specific education conditionalities, and because of age-specific drop out rates. Indeed, the CCT's incentives to continue schooling are likely to affect schooling behaviour among those age groups that are at the highest risk of drop out. In contrast, the education component is likely to have a limited impact for children who were likely to have been in school in the absence of the programme. Our analysis therefore identifies a cohort of boys and a different cohort of girls based on school drop out age patterns in order to establish gains in educational attainment, learning and labour market outcomes.

The theory of change indicated above provides the conceptual base for interpreting the differential effects. In particular, we interpret these differentials as mainly capturing differential impact due to differences in age of exposure. In order to interpret them as such, it is important to discuss the similarities in the programme and the context between the early and late treatment group.

In most ways, the RPS design was constant over time, though there were some changes during the phasing in of the late treatment group. In particular, the absolute value of transfers for the programme's health component was reduced by 20 per cent on average, while the education transfer remained about the same. There was also continued supply of health services, without conditionality, to the early treatment group even after the end of the transfers. More substantively, between 2003 and 2005, young teenagers aged 12–14 from beneficiary households were required to attend information

[4] Increased labour market returns could also come from other components of the intervention.

sessions on reproductive health and contraception, as well as the *titular.* Healthcare providers also made contraceptives available to beneficiaries. We analysed the potential consequences of this change in our qualitative work in Section 5.3.

Even if there had been no critical changes in the education, nutrition and health components of RPS over time, nor the introduction in these areas of other related programmes, it could still be possible that differences in programme effect during the early treatment and late treatment periods were due to other contextual differences across the calendar years over which the programme operated. For example, world coffee prices declined substantially in 2001 and were at their lowest level in decades in 2002, recovering slowly after that. Purchasing power parity adjusted real GDP per capita growth declined from 2.4 per cent in 2000 to –0.6 per cent in 2002, recovering thereafter (World Bank 2013). Households in the late treatment therefore experienced declines in expenditure prior to receiving the programme.

Despite these differences over time, however, many other contextual factors were fairly consistent. The same political party was in power throughout the life of the programme. No significant natural disasters occurred during the period and the programme finished before the global financial crisis. Moreover, experimental comparison of early treatment versus control in 2002, and late treatment versus early treatment in 2004 indicates that RPS had effects of similar size on expenditure (Maluccio 2010). On balance, then, while it is not possible to claim the early and late treatment periods were identical, it is plausible that differences between them do not drive the results.

### 3.4 Primary outcomes to be examined

Given that we study the impact of RPS 10 years after it began, the primary outcomes of this evaluation are, as much as possible, those related to long-term impacts – cognitive and achievement test scores, anthropometric measures, labour market outcomes and household expenditures in particular. These measures rely on the two main instruments used for the 2010 follow-up: a household-level survey with detailed information on labour markets and an individual-level survey with a battery of cognitive and achievement tests, as well as anthropometric measures. See Appendix F for a description of the tests. To shed light on the theory of change and the underlying mechanisms, we also use the earlier evaluation surveys to establish how short-term impacts on both schooling and anthropometric measures are related to the final outcomes. Information from a qualitative survey provides further information on the theory of change and mechanisms – for example, on the sustainability of impacts on expenditure behaviour.[5]

## 4. Tracking of programme beneficiaries

In contrast to the majority of evaluations 3ie has funded recently, rather than a real-time evaluation, this research project was a follow-up to an intervention completed five years before the present research project began. Section 3 gives a brief description of the intervention. Instead, we use this section to detail the careful migrant tracking procedures we used, and their results, which are likely to be of interest to the evaluation community.

[5] For the qualitative analyses, we draw on semi-structured interviews in a dozen early treatment, late treatment and comparison communities.

Before providing a detailed summary, we outline the importance of tracking migrants for the overall aim of the study. Given the research questions and the potential (but unknown) relationship between cognitive development, education and migration, it was paramount to our study to minimise attrition. This was not only important to maximise statistical power, but also to enhance the internal and external validity of our research. As households and individuals that migrate are likely to have different characteristics from those who stay (as seen in the case of *Oportunidades*, Parker *et al.* 2008), the sample we would have surveyed had we not tracked migrants is unlikely to be representative of the entire initial target population. This has clear implications for external validity. Moreover, migration might itself have been affected by treatment, and the characteristics of those who migrate or attrite in the different treatment groups might be different, potentially leading to substantial biases in the estimates and affecting internal validity. While these concerns are relevant for all evaluations, they are particularly pertinent for our study given the long period (10 years) and the focus on young individuals who tend to be particularly mobile (and hence have high potential attrition). Extensive tracking procedures were therefore a key component of our research design: we targeted all households (of the original *titular*) for follow-up, and all individuals under 22.

Of the 3,503 households from 12 municipalities in the targeted sample, 21 per cent could not be interviewed in their original location. We tracked all households and individuals that could not be found in their original locations to their new locations in Nicaragua and Costa Rica (the destination of over 95 per cent of international migrants from the sample). As migration is often temporary, we organised multiple return visits to the original locations for the field team to incorporate temporary migrants who might have returned. As a result, final attrition at household level is less than 7 per cent. Importantly, there is no significant difference in attrition between the experimental early and late treatment groups, nor between the experimental late treatment group and the non-experimental comparison group. There were also few substantial differences between those found and those not found in the 2010 follow-up.

At the individual level, of the 10,977 individuals under 22 who were targeted to be included in the sample for the individual survey, 41 per cent could not be interviewed in their original location. Of those, approximately 19 per cent were temporarily absent, while the remaining 22 per cent had migrated to other households, often in other locations in Nicaragua or Costa Rica. As a result of extensive tracking, final individual attrition due to permanent migration for those aged less than 22 is 9 per cent. For 5 per cent, however, we have proxy information on the individual from the household survey. As with household-level attrition, there are no significant differences in attrition among early treatment, late treatment and non-experimental comparison groups for individual attrition in the household survey, nor for attrition in the individual survey.[6]

These rates of attrition compare favourably with other impact evaluation studies covering similar periods and populations, and with many longitudinal studies that cover much shorter periods. In contrast to many other longer-term studies, we tracked all households and targeted individuals, rather than a random subsample, to both increase statistical power and better capture heterogeneity that might be related to different destinations.

[6] The balance between the different groups for the cohort relevant for the results in this report is discussed in Section 5.

Tracking was done in four stages. First, during the regular survey period carried out in all 12 original municipalities, survey teams tracked and immediately included all households and individuals that had moved to another location in the same or a nearby community. In cases of temporary absence, the team revisited the households in subsequent days as long as they were working in the area. Location information for everyone else was recorded at this time. This regular survey period lasted approximately four months.

Second, after completing these first-round visits, tracking teams returned to each of the 12 study municipalities to find and survey many of the individuals and households who had been away during the first round. During this second round, at least two different teams visited each of the original communities. We also tracked households and individuals who had moved to other locations within the original 12 municipalities. We verified previously collected information on the destination of migrants. Concurrently, a separate team was tracking in Managua and its surroundings (the dominant migrant destination) and teams communicated updated information on migrant movements in real time. This second round lasted approximately four months.

In the third round of tracking, we extended the search to all destinations in Nicaragua; it lasted approximately seven months, which included a three-month cessation of all field activities during the rainy (and hurricane) season when road access in many rural areas was very difficult and potentially dangerous. During this phase, field teams were simultaneously operating in different parts of the country and continued communicating all updated information on migrant movements. Lastly, we undertook additional return visits to the original locations during the Christmas and New Year break, as many migrants (temporarily) return to visit their family at that time of year.

Finally, after obtaining additional funding,[7] we expanded tracking to Costa Rica, the destination of nearly all international migrants. This final round lasted approximately two months.

## 5. Impact results

In this section, we describe our research findings to date, organised according to four principal research questions.

### 5.1 Early child development and critical ages for intervention

RPS purposively targeted highly impoverished rural areas where malnutrition for children under five was severe (40 per cent) and large cognitive deficits were likely. While the short-term evaluations did not collect information on cognitive outcomes, the data we collected as part of the long-term evaluation confirms that large cognitive deficits indeed characterise the study population (Table D1, Appendix D). In particular, there are very substantial delays in various domains of early childhood development for children aged six months to six years, as assessed by the Denver Developmental Screening Test.[8]

[7] National Science Foundation SES grants 11239945 and 1123993.

[8] We assessed children's social-personal, language, fine motor and gross motor skills using the four sub-scales of the Denver Developmental Screening Test (Frankenberg and Dodds 1996). As

The largest delays are in language development, with 84 per cent of children in the bottom quartile of the international distribution, and 63 per cent in the bottom decile. Delays are also common for social-personal and fine motor development, but somewhat less for gross motor development. Overall, the magnitude and comparison across domains is very similar to those found for a nearby population in 2006 (Macours *et al*. 2012). In addition, delays increase with age for personal-social and language development, while they decrease for gross motor development. For language development, where delays are the largest and increasing by age, 50 per cent of children are already delayed in their first three years of life.

With this context, we turn to our key impact findings in the area of early child development.[9] Previous analysis[10] has demonstrated short-term impacts on nutrition and health for young children in the early treatment group. Given these earlier findings, and the programme's subsequent phase-in to the late treatment group, this is a good setting to test the hypothesis that the first 1,000 days are the critical window for both cognitive skill formation and physical development. To maximise power and explore heterogeneity, we focus on boys first, since they are more vulnerable in early life than girls, particularly during the pre-natal period (Eriksson *et al.* 2010). Using the 2010 data we estimate intent to treat effects on differential cognitive outcomes (measured by seven different cognitive tests) and physical growth (measured by height-for-age standardised Z-score and weight-for-age standardised Z score) of boys now 10 years old who were exposed to the programme starting *in utero* and up to age two in the early treatment group, versus those exposed between the ages of two and five in the late treatment group.

For the average of all cognitive tests, outcomes for boys exposed to the programme *in utero* are a statistically significant 0.15 standard deviation (sd) larger than boys exposed later (Table 1, Panel A). These gains are fairly similar or higher across the different tests administered, which included tests on memory, processing speed and visual integration. Results are shown in Table 2.

In contrast, however, for boys' anthropometrics, there is an insignificant –0.07 sd differential effect (see Table 1). Hence, receiving treatment starting *in utero* and for the first two years of life did not lead to higher physical growth 10 years later, when compared to receiving treatment later. In conjunction with the short-term results on growth, this zero differential effect suggests catch-up growth in the late treatment group. This interpretation is supported by a new analysis of the programme's administrative data, which also showed substantial gains in physical growth in the short run, and indications of catch-up even after only one year of programme exposure in the late treatment group.[11]

Examining 10-year-old girls in early versus late treatment groups using the same methodology, we find no differential effect of early versus late treatment (see Table 1, Panel B). The lack of differential results between early and late treatment for girls does

---

the Children's Learning Achievement Profile, widely used in Nicaragua for child monitoring, is a slightly modified version of the Denver Scale, the test is appropriate for the population we study.

[9] This section draws heavily from Barham *et al.* (2013a).
[10] See Section 2.1 and Barham *et al.* (2013a).
[11] See Figure 1 in Barham *et al.* (2013a).

not reflect differences in the programme or in take-up at household level, by gender. Remember, however, that short-term results did show substantial gains in nutrition and health of young children, including girls. So girls did benefit, but they did so equally in the early and late treatment groups, leading to no differential effect. In contrast for boys, benefitting in the 1,000-day window is more important, consistent with other evidence of boys being particularly vulnerable to very early life circumstances.

### 5.2 Schooling, learning and the labour market

Consistent with the early childhood development outcomes, the highly impoverished areas targeted by RPS had low levels of schooling. For example, enrolment rates for children targeted by the programme (aged 7–13, not yet having completed fourth grade) were less than 70 per cent in 2000, reflecting both late-for-age starts and early drop outs. Since boys and girls had differential drop out rates by age, we expected heterogeneous effects and therefore examined them separately.

For the differential impact assessment for boys, we focus on ages 9–12, as this cohort is the most likely to have benefitted more intensively from the education transfers in the early treatment group than the same age children in the late treatment group, as many of the children in the late treatment group would have been too old to be eligible for the education transfers by the time their communities were phased into the programme. Furthermore, for boys, this age cohort encompasses the ages where the risk of school drop out, in the absence of the programme, was high, further increasing the potential impact of the programme (see Figure 1).

Short-term estimates using the original evaluation data confirm previous research, indicating that by 2002 the programme had led to a half-year increase in the grades attained and by 2004 an increase of 0.7 years (See Table 1 in Barham *et al.* 2013b). This indicates that, at least by 2004, the programme had led to a sustained increase in the number of grades attained for the early treatment group.

By 2010, seven years after the early treatment group stopped receiving the transfers, boys (now young men aged 19–22) in the early treatment group had still attained a half-year more grades than those in the late treatment group (see Table 3, Panel A). Moreover, the increase in grades attained was accompanied by gains in learning, in contrast to the findings for *Oportunidades* where a similar sized differential increase in grades attained was not coincident with gains in learning (Behrman *et al.* 2011).[12] In particular, individuals in the early treatment group saw an average improvement of a quarter of a standard deviation on standardised tests in maths (speed and problem solving) and Spanish (reading and spelling). Non-experimental estimates of the absolute impacts, obtained by comparing the late treatment to the comparison group in 2010 through matching, show absolute effects of more than half a year of education attained and improvements in learning of between 0.15 and 0.25 sd on most of the tests, though statistical significance varies depending on the estimator (see Appendix G for more details on the non-experimental methods).

In contrast with the results for maths and language tests, we found no significant impact on the Raven test, a widely used measure for cognitive ability, as opposed to skills

[12] See Barham *et al.* (2013b) for a more detailed discussion of learning results and identification strategy.

learned in school (Raven *et al.* 1984; Table 3 in Barham *et al.* 2013b). This is consistent with cognitive development taking place mostly during early childhood before the programme had begun for these individuals.

Did the programme yield dividends in the labour market for these young men? To answer this question, it is useful to start from this cohort's overall labour participation in the context of a poor rural area of Nicaragua. Almost all (98 per cent) of the young men are economically active; more than 80 per cent work in agriculture on the household farm which is typically still managed by their father; and more than 80 per cent also work outside the home in various forms of wage labour.

Figure 3 shows the individual earnings per month worked and total individual earnings from all activities outside the household farm in the last 12 months (regardless of whether they were working or not).[13] We see a clear rightward shift in the distributions of both monthly and total earnings for young men in the early treatment group, indicating a general and significant increase in labour market returns. The early and late treatment group distributions are significantly different in both cases, providing evidence of positive labour market returns. Similar results are found in linear regressions of the average differential effect of the programme.

To understand the mechanism driving these differences, the first important finding is that those in the early treatment group are approximately 4 per cent more likely to have worked outside the household farm in the previous 12 months (see Table 4). We also find that early treatment young men are approximately 8 per cent more likely to have migrated temporarily for work, spending on average an additional 10 days away from home. This increased labour force participation and seasonal migration did not represent a shift to jobs requiring substantial higher education levels; jobs are primarily concentrated in non-skilled wage work, including agricultural day labour activities. This is in contrast to findings from *Oportunidades,* where there is some evidence of a shift from agricultural to non-agricultural labour (Behrman *et al.* 2011).

To understand these results and the related policy implications, the magnitude of the learning effect and the context are important. Prior to the programme, the average grade attained was less than Grade 4. We find increases for certain groups of approximately half a grade. Consequently, it would have been illogical to expect such gains would result in young men massively shifting to high-skilled wage jobs. But the gains in learning did translate into income gains, and hence point to real sustained benefits from receiving the CCT programme during primary school years. A plausible interpretation of these results is that the young men with a little bit more education (due to the intervention) become better at finding higher-paying jobs that are often further away from home. Increased maths skills might help them make the right cost-benefit calculations of temporary migration, and literacy skills might help them complete the necessary paperwork and obtain temporary jobs abroad.

[13] Approximately 89 per cent of the young men do not yet head their own household; agricultural self-employment mainly involves work on their father's farm, making it difficult to calculate individual returns for those activities.

Hence even in a context where work in good non-agriculture jobs is difficult to find, increased education due to the programme could result in higher labour market returns.[14]

Our assessment of the impact of RPS on schooling, learning and labour market outcomes for boys hinged importantly on the timing of the programme transfers and conditions relative to when boys normally dropped out of school (See Figure 1 in Barham *et al.* 2013b). These patterns were different for girls, however, as they tended to drop out of school two to three years later than boys (see Figure 2). Consequently, in analysing them we extended the critical age range to include girls who were 13 or 14 at the start of the programme. Although these girls were not, in principal, subject to the programme conditionalities, transfers to their households (in which there may have been other children who were eligible and subject to conditionalities) could still lead to increases in schooling. Examining the girls in two-year intervals, we find: a small (0.10) insignificant increase in grades attained for 9–10-year-olds (aged 19–20 when interviewed in 2010); no effect on 11–12-year-olds; and a substantial and significant effect of three quarters of a grade on 13–14-year-olds. The increase for the older girls is remarkable and indicates a potential important spillover of the programme on older girls who were not eligible for the education transfers. We interpret the findings for the 11–12-year-old girls as evidence of programme impacts on both the early and the late treatment group, leading to no significant differential impact in grade attained. Consistent with the small differential effects on grades attained, we find no differential effects on learning for 9–12-year-old girls. Taken together with the boys, this strengthens the interpretation that the learning effects are the result of additional grades attained. Unfortunately, cost considerations prohibited us from administering achievement tests to the older girls, so we are unable to assess whether they experienced learning gains.

### 5.3 Reproductive behaviour

In this section, we draw on qualitative work completed prior to the quantitative survey to explore the relationship between the programme and reproductive behaviour. More fully described in Appendix E, the objectives of the qualitative study were to understand how the programme operated and its effects from the point of view of the beneficiaries, persons involved in programme execution and others directly or indirectly involved in the programme, such as teachers and health workers. Another objective was to explore why and by what mechanisms within households or communities any effects may have occurred.

The principal methods we used in the study were semi-structured interviews with individuals and in focus groups, including beneficiaries (particularly mothers of children benefiting directly from the programme), voluntary *promotoras* (promoters) who worked with the programme, teachers, health workers, leaders and others involved in the programme's execution. Dividing the 12 municipalities into four areas, we interviewed three communities in each area, one each from the early treatment, late treatment and non-experimental comparison group.

[14] In the context of poor rural Nicaragua, seasonal migration comes with important income gains, consistent with other recent evidence on the potential returns of lifting constraints to seasonal migration (Bryan *et al.* 2013) and with findings in Macours and Vakis (2010) on positive impacts of seasonal migration in another region of Nicaragua. To understand the broader welfare implications of such migration, one might also want to account for the non-monetary costs, including the socio-emotional costs of family separation, but this goes beyond the scope of the current study. Instead, we refer to the larger literature on potential welfare trade-offs of migration (Stillman *et al.* 2012).

During the first years of the programme, one of the first health education workshops for the main early treatment beneficiaries focused on reproductive health and contraception. When the control group was phased in, this was expanded to young teenagers from beneficiary households (age 12–14), who were also required to attend information sessions on reproductive health and contraception. In this later period, healthcare providers also made contraceptives available to all beneficiaries.

The sessions for adults and for teenagers were the first of their kind in nearly all of the communities. The adolescent workshops, in particular, appear to have been very salient, leading to animated discussions in several of the qualitative interviews more than four years after the last of them took place. The study revealed apparent increases in knowledge and use of contraceptives in most treatment communities during the programme. Some commented that these patterns translated into reductions in teenage pregnancies. What was less clear is how well family planning, and in particular the supply of contraceptives, was maintained. Changed attitudes, knowledge and practice following the educational sessions may not have extended to continued practice with inconstancy of supply after the programme ended.

### 5.4 Indicators of sustainability

The evidence presented above has shown important gains – for example, in cognition of school-age children, and in learning and labour market outcomes for young adults. These are clear indications of the programme's persistent or sustainable effects. In this section, we turn to some other potential indicators of sustainability – in particular, continued investment in human capital of young children – and more general household expenditure patterns. For the former, we use the qualitative study and for the latter, the double-difference non-experimental matching techniques that incorporate the comparison group into the analysis.

An overarching question in the qualitative research was to explore, from the beneficiaries' perspective, whether the increased investment in their children's nutrition, health and schooling continued after programme ended, and why or why not. Echoing earlier qualitative findings from a study carried out during the intervention (Adato and Roopnaraine 2004), it was clear that beneficiaries recognised the importance of such investments and valued them. It is even likely that attitudes shifted due to the programme, though we cannot ascertain that with confidence.

Less clear, however, is whether such investment could continue after the programme for these or other children – for two primary reasons: the cessation of transfers and uncertain supply-side considerations. For example, several beneficiaries indicated that, while they would want to continue sending their children for further schooling, without the transfers it was less feasible. For others, an additional problem was access to schooling, particularly for those progressing on to secondary level. Similar constraints were expressed with respect to nutrition and health considerations for younger children, in particular, the inability to purchase some of the recommended highly nutritious foods and the cessation of the direct programme-provided health services. From this angle, then, it is clear that there are potential constraints to the sustainability of programme impacts through behavioural changes.

An alternative approach to investigating sustainability is to assess whether, several years after the end of the programme, participating households had higher levels of economic well-being, as measured by per capita expenditures or food shares (percentage of expenditure spent on food). We assessed this using non-experimental matching methods, that first were internally validated using the short-term evaluation (Barham *et al.* 2012).[15]

We present non-experimental results for a number of matching estimators in Table 5. These results compare 2010 households from the late treatment group with those from the comparison group (see Appendix G for more details on non-experimental methods). The intention to treat effects for log per capita expenditures are positive and suggest that expenditures are 7–11 per cent higher in the late treatment group than the comparison group. However, none of the estimates are statistically significant, at the 5 per cent level or higher. Similarly, there is no significant effect on food shares. Taken together, four years after households stopped receiving transfers, there is little evidence of continued higher food-related per capita expenditure or expenditure shares among the original programme households.

### 5.5 Internal and external validity

The internal validity of the results on early childhood critical ages (Section 5.1) and on schooling, learning and labour market outcomes (Section 5.2) is based on the experimental variation in the data. As the experiment resulted in a balanced sample at baseline, and as the intensive tracking was successful in maintaining such balance even after 10 years, the internal validity of this set of results is strong. This is also illustrated by Table D2 (Appendix D), which shows the baseline balance for both the tracked and non-tracked boys in the 9–12 age group.

The internal validity of the results based on the non-experimental matching method estimations (Appendix G) depends on the plausibility of the identifying assumptions. To shed light on this, we have attempted to validate the matching estimations by comparing experimental estimations with the matching estimations using the short-term evaluation surveys (Barham *et al.* 2012). In addition, we test the robustness in various ways. Overall, we believe this provides good support for the identifying assumptions.

To understand the external validity of the findings, a first key consideration is the comparability of the RPS programme with other CCTs. RPS was in fact modelled after

[15] Experimental methods comparing the differential effect for these two household-level variables between the early and late treatment households do not provide a good estimate of sustainability effects because both groups received the programme for three years. At the household level, the differential effect only estimates the effect of receiving transfers three years earlier, which conceptually could lead to higher or lower household expenditure after the end of the programme. Expenditure could be higher if there are multiplier effects of possible investments that households made with the transfers, lower if there is fade-out of behavioural change over time, or the same if neither of these are important or if they cancel each other out.
Experimental evidence presented in Table 5 shows there are no statistically significant programme effects on log per capita expenditures or food shares. This may seem to contrast findings for Mexico, where there was a differential effect three years after the control was phased in (Gertler *et al.* 2012). The comparison is, however, imperfect: the Mexican programme continued after the phase-in, while for RPS both experimental groups had long stopped receiving transfers in 2010.

*Progresa/Oportunidades* in Mexico, and indeed most of its key design features are very similar to other CCTs in Latin America. This is important for the broad relevance of these findings. That said, the programme is different from others in some key design features –notably in its limited length (three years) and focus on only the first four years of primary school. In addition, the rural Nicaraguan context in which it was implemented is poorer than much of the rest of Latin America, and baseline levels of education lower. Therefore, it might have been easier to obtain gains in learning and young adult labour market outcomes in this context than it would be in other countries, where children typically reach higher levels of schooling. It is possible, therefore, that the external validity of this study might be greater for other high-poverty settings, such as Sub-Saharan Africa or South Asia.

Given these observations on external validity, and the fact that we offer longer-term and more final outcomes, the findings of this study are arguably particularly important for cost-effectiveness calculations of CCT programmes.

## 6. Conclusions and policy recommendations

CCT programmes have become the anti-poverty programme of choice in many developing countries. Their approach – combining short-term poverty reduction with enhanced investment in human capital – has widespread policy appeal. A number of rigorous empirical studies have established that these programmes are effective at reducing short-term poverty and increasing health and children's school attainment. There is little evidence, however, on whether those short-term benefits translate into the longer-term gains needed to fully justify these programmes. As the policy discussion on CCTs shifts towards designing exit strategies, there is a demand from policymakers for establishing whether such longer-term gains exist.

This research provides rigorous evidence demonstrating such longer-term gains and providing key inputs into this policy debate. Based on our findings of substantial long-term gains on early childhood cognitive development and educational achievement, we conclude that continued support for this type of programme is warranted. In addition, our research allows deriving more nuanced policy recommendations regarding the optimal timing of the different components of CCT interventions. Moreover, the lessons about timing are more widely applicable to other types of interventions targeting early childhood or primary school, and underscores that these may differ by gender.

In particular, our results demonstrate that boys exposed *in utero* and during the first two years of life have better cognitive outcomes when they are 10 years old than those exposed in their second year of life or later. These results confirm that interventions that improve nutrition and/or health during the first 1,000 days of life can have lasting positive impacts on children's cognitive development. However, there are no differential impacts on anthropometrics, despite short-term differences resulting from the programme, demonstrating that complete catch-up in anthropometrics was possible. We also find catch-up in physical growth for girls. These results are consistent with other empirical evidence of catch-up for physical growth as well as with the medical literature on brain development.

For school-age children, our research shows important long-term effects on educational attainment and learning. In particular, we estimate statistically significant and

substantial average achievement gains in maths and language of a quarter of a standard deviation. As these estimates are obtained for children from households that stopped receiving transfer seven years earlier, they provide unique evidence on the sustainability of CCT impacts. The findings on learning are important as they point to the potential for other longer-term gains. Indeed, they are a potential important mechanism underlying the positive impacts on labour market outcomes for these same individuals. Further research aims at a more comprehensive understanding of the longer-term labour market outcomes.

The findings on sustained learning outcomes are also important in their own right, and critical from a policy perspective. In particular, they establish that gains in learning can be achieved in poor rural settings, despite the many other constraints on schooling quality that likely still existed. Recently, many questions have been raised about the potential of schooling programmes in general, and of CCTs in particular, to affect learning outcomes (see, for example, Filmer *et al.*; Levy and Schady 2013). Our research provides a clear positive answer to this question, an important conclusion for policymakers. Indeed, while complementary interventions, such as targeting schooling quality, might be useful to further enhance children's learning, our research indicates that CCT programmes can be an important part of the solution in high-poverty settings. And similarly, even with a variety of (supply) constraints in the labour market, there can also be important labour market returns, despite relatively low levels of education.

Overall, the results clearly show that CCTs can lead to sustainable gains in human capital accumulation, and that these remain long after households exit the programme. This suggests their potential to lead to long-term poverty reduction.

# 7. Tables and figures

Table 1: 2010 Intention to treat differential effects on cognition and anthropometrics

| | Born in first 12 months of programme | | Born in first 6 months of programme | |
|---|---|---|---|---|
| | Cognition | Anthropometrics | Cognition | Anthropometrics |
| PANEL A: BOYS | | | | |
| *Age and strata controls* | | | | |
| T (early treatment) | 0.147** | −0.068 | 0.173*** | −0.086 |
| | (0.060) | (0.100) | (0.063) | (0.118) |
| *All controls* | | | | |
| T (early treatment) | 0.145** | −0.096 | 0.155** | −0.146 |
| | (0.062) | (0.096) | (0.069) | (0.101) |
| Observations | 368 | 368 | 267 | 267 |
| PANEL B: GIRLS | | | | |
| *Age and strata controls* | | | | |
| T (early treatment) | 0.013 | 0.055 | −0.022 | −0.050 |
| | (0.077) | (0.098) | (0.082) | (0.103) |
| *All controls* | | | | |
| T (early treatment) | 0.021 | −0.006 | −0.020 | −0.082 |
| | (0.072) | (0.077) | (0.086) | (0.088) |
| Observations | 340 | 340 | 246 | 246 |

Notes: ***p < 0.01, **p < 0.05, *p < 0.1.
Standard errors are clustered at the locality level and in parentheses.
SURE estimates of seven cognitive and two anthropometric measures all expressed as Z-scores, following Kling *et al.* (2007).
Controls include: monthly age, strata and test administrator dummies, mother's education and birth order, baseline household size, living standard (estimated proxy means), assets (number of rooms, land, radio, animals, tools), distance to school and locality means for anthropometrics of under three-year-olds. See Barham *et al.* (2013a).

Table 2: 2010 Intention to treat differential effects on cognitive test scores, boys

| | Cognition | | | | | | | Anthropometrics | |
|---|---|---|---|---|---|---|---|---|---|
| | Short-term memory | Long-term memory | Raven | Process speed | Receptive vocabulary | Forward digit span | Backward digit span | Height | Weight |
| *T* | 0.156* | 0.0542 | 0.229** | 0.228** | 0.229** | 0.155* | –0.0364 | –0.0629 | –0.129 |
| | (0.0912) | (0.115) | (0.109) | (0.0948) | (0.0931) | (0.0857) | (0.114) | (0.100) | (0.122) |

Notes: ***p < 0.01, **p < 0.05, *p < 0.1.
N=371.
Boys born in the first 12 months of the programme.
Standard errors are clustered at the locality level and in parentheses.
All outcomes expressed as Z-scores.
Controls include: monthly age, strata and test administrator dummies, mother's education and birth order, baseline household size, living standard (estimated proxy means), assets (number of rooms, land, radio, animals, tools), distance to school and locality means for anthropometrics of children under three years old.

Table 3: 2010 Intention to treat and matching results on schooling outcomes and test scores, boys

| | Grades attained | | Standardised test scores (Z-score) | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Absolute | Z-score | Reading fluency | Spelling | Word identification | Maths fluency | Maths problems | Receptive vocabulary |
| PANEL A: Experimental results for boys aged 9–12 in 2000 at start of programme | | | | | | | | |
| ITT | 0.501* | 0.159* | 0.320*** | 0.252*** | 0.228*** | 0.259*** | 0.175** | 0.176** |
| | (0.274) | (0.087) | (0.082) | (0.077) | (0.075) | (0.082) | (0.065) | (0.078) |
| PANEL B: Matching results for boys aged 9–12 in 2003 when late treatment group started to receive transfers | | | | | | | | |
| NN1 | 0.597 | 0.186 | 0.150 | 0.246** | 0.189* | –0.002 | 0.247+ | 0.028 |
| | (0.737) | (0.235) | (0.150) | (0.084) | (0.091) | (0.140) | (0.138) | (0.168) |
| NN5 | 0.909* | 0.283* | 0.184+ | 0.223** | 0.207** | 0.092 | 0.224* | 0.096 |
| | (0.393) | (0.124) | (0.110) | (0.048) | (0.018) | (0.077) | (0.107) | (0.124) |
| Kernel | 0.871+ | 0.271* | 0.158 | 0.207* | 0.198 | 0.077 | 0.212* | 0.044 |
| | (0.479) | (0.134) | (0.132) | (0.087) | (0.125) | (0.117) | (0.097) | (0.125) |
| LLR | 0.822* | 0.256+ | 0.132 | 0.194* | 0.173 | 0.074 | 0.228* | 0.040 |
| | (0.390) | (0.151) | (0.141) | (0.093) | (0.124) | (0.135) | (0.107) | (0.138) |
| Observations | 359 | 359 | 315 | 322 | 321 | 320 | 321 | 322 |

Notes: ***p < 0.01, **p < 0.05, *p < 0.1.
Panel A includes dummies for stratification groups and three monthly age dummies, and displays the experimental results that provide the differential effect between the early and late treatment groups.
Panel B displays the estimated absolute effects from matching the late treatment group to the comparison group.
Results for four matching models are shown: nearest neighbour using 1 and 5 neighbours (NN1 and NN5), kernel matching and local linear regression (LLR) matching. Nearest neighbour specifications control for 2002 values of the outcome.
Standard errors are clustered at the locality level and are given in parentheses, except for kernel and LLR models where they are bootstrapped.
For more details on Panel A results, see Barham *et al.* 2013b, Table 2.

Table 4: 2010 Intention to treat differential effects on labour market outcomes for young men (boys aged 9–12 at start of the programme)

| | Works outside household farm | Participation in any economic activity during temporary migration | Months in any economic activity during temporary migration | Months in non-skilled wage employment during temporary migration |
|---|---|---|---|---|
| ITT | 0.0397* | 0.0786*** | 0.304** | 0.345*** |
| | (0.023) | (0.026) | (0.130) | (0.121) |
| Mean in control | 0.829 | 0.317 | 0.995 | 0.840 |

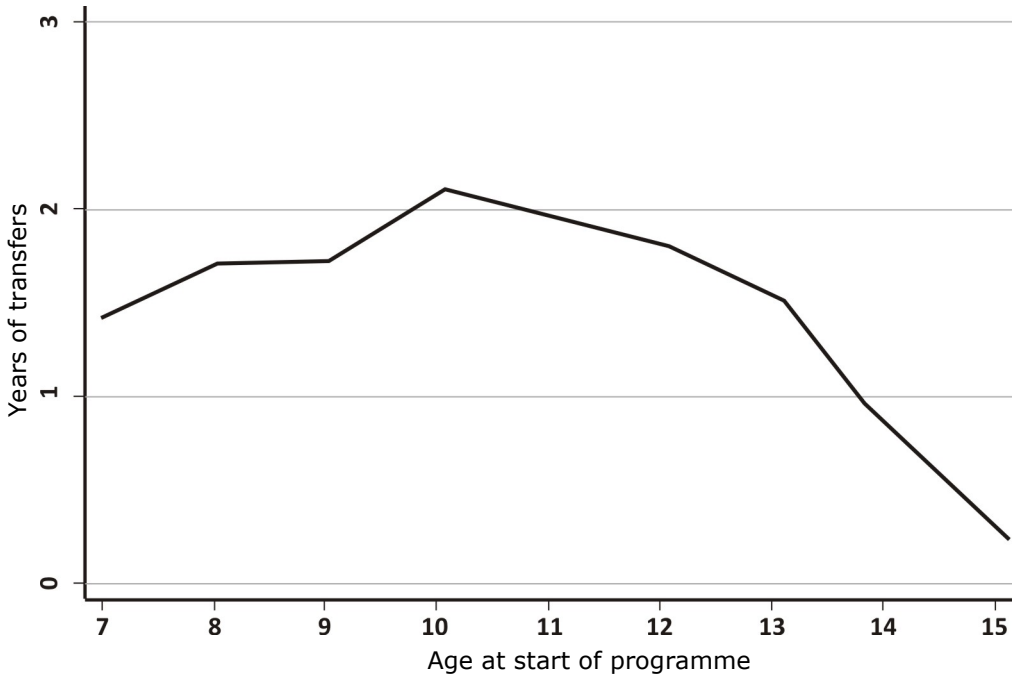
Notes: ***p < 0.01, **p < 0.05, *p < 0.1.
N=371.
Standard errors are clustered at the locality level and are in parentheses.
Controls include: three-month age dummies, strata and regional fixed effects, log of *per capita* expenditures (as estimated by the proxy means) in 2000, distance to school, grade attained and an indicator variables for whether in 2000: the household was active in agriculture, the child had no completed grades, the mother and father were in the household and the respondent was the child of the household head.

Table 5: Intention to treat differential and matching results for per capita expenditure and food shares

| Outcome | Experimental (intention to treat) | Matching | | | |
|---|---|---|---|---|---|
| | | NN1 | NN5 | Kernel | LLR |
| Log per capita Expenditures | −0.056<br>(0.045) | 0.018<br>(0.032) | 0.072*<br>(0.031) | 0.096<br>(0.068) | 0.106<br>(0.068) |
| Food share (% of total expenditure) | 0.008<br>(0.011) | 0.002<br>(0.019) | −0.006<br>(0.014) | −0.017<br>(0.017) | −0.013<br>(0.020) |
| Observations | 1,392 | 744 | 744 | 744 | 744 |

Notes: ***p < 0.01, **p < 0.05, *p < 0.1.
ITT differential refers to the differential effect between early and late treatment for the experimental models, and to the absolute effect between the late treatment and comparison group for the matching models.
Results for four matching models are displayed: nearest neighbour using 1 and 5 neighbours (NN1 and NN5), kernel matching and LLR matching.
All specifications control for 2000 values of the outcome except for kernel and LLR matching models.
Standard errors are clustered at the community level for models except kernel and LLR.
Bootstrapping is used to determine standard errors for the kernel and LLR models.

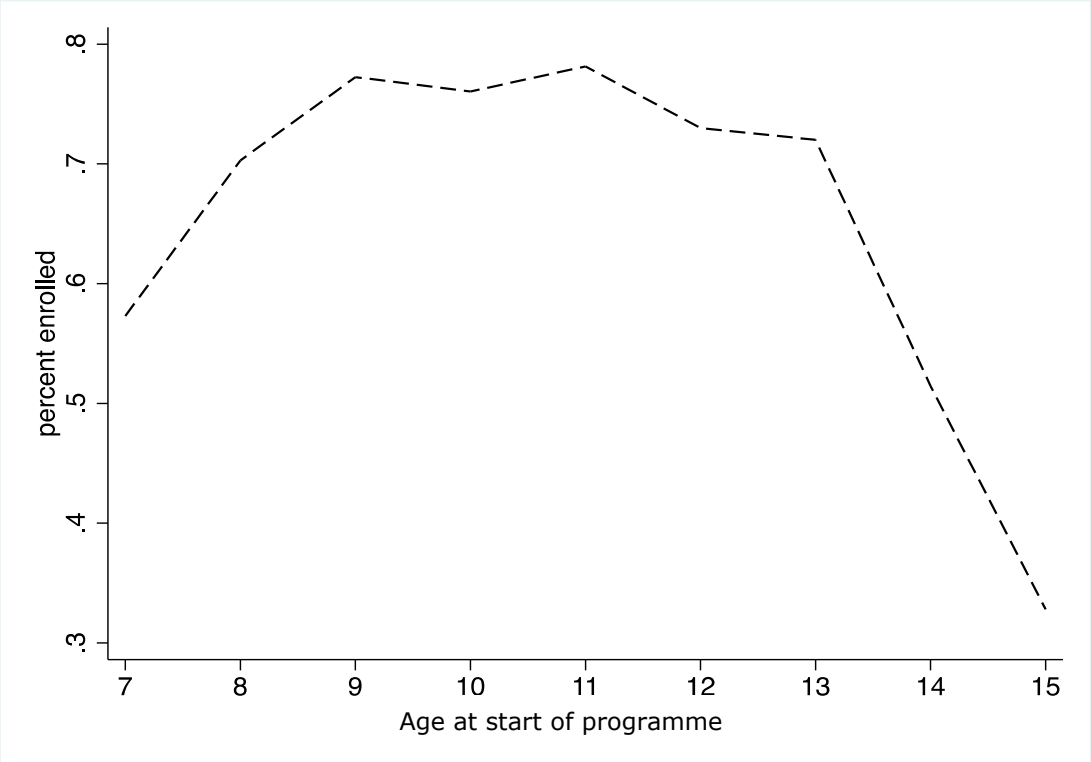
Figure 1: Percentage of boys in school in 2000 by age




Source: Authors' calculations using 2000 census data collected for the RPS project.

Note: Age at start of programme refers to the person's age in November 2000.

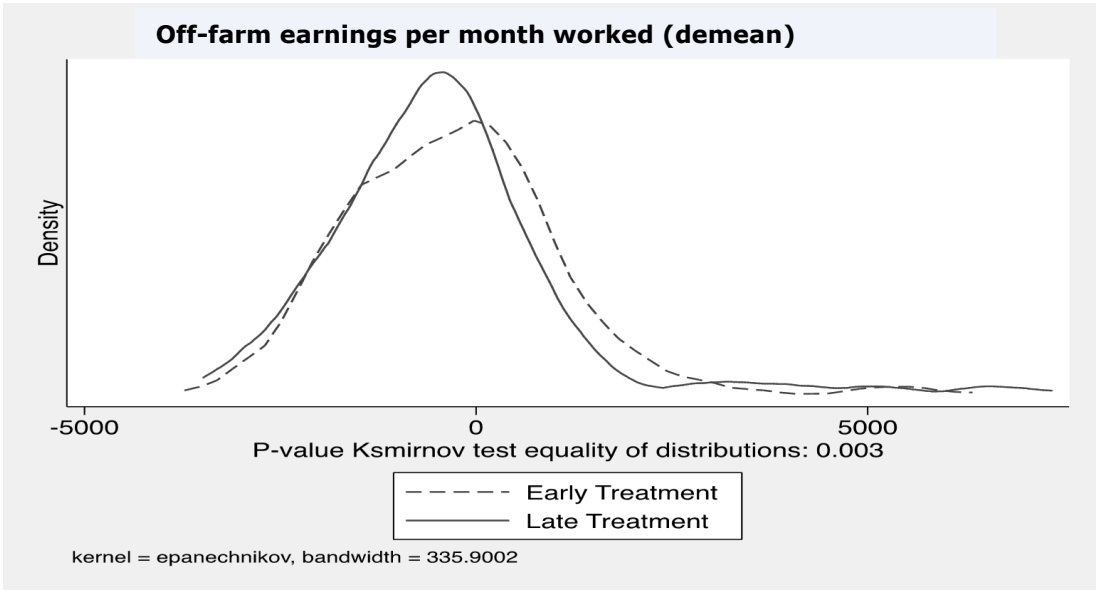
Figure 2: Percentage of girls in school in 2000 by age

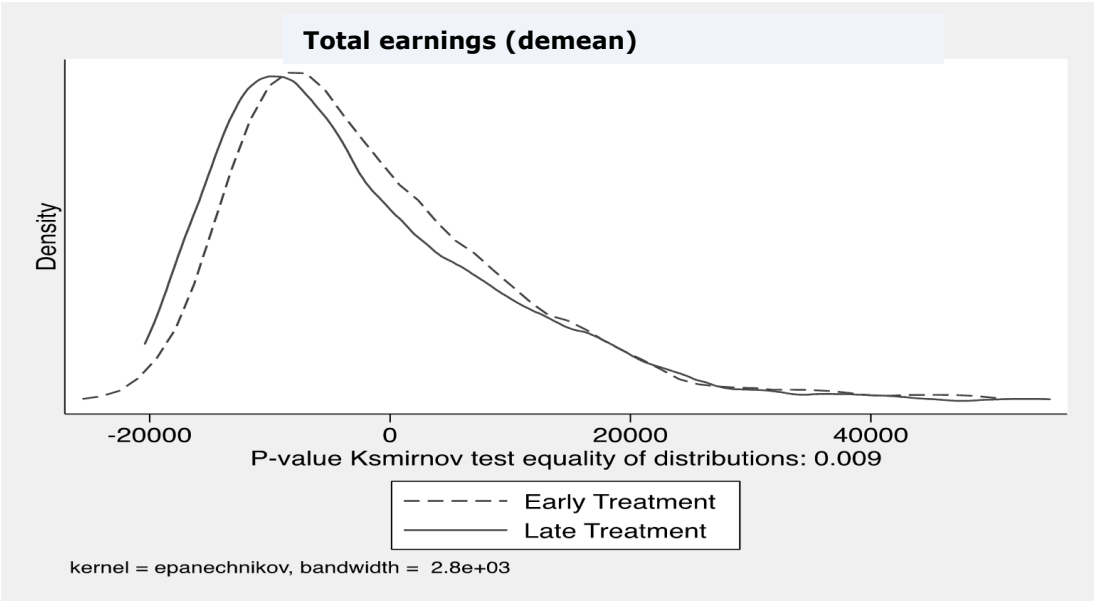



Source: Authors' calculations using 2000 census data collected for the RPS project.

Note: Age at start of programme refers to the person's age in November 2000.

Figure 3: Differential effect on earnings for boys aged 9–12 at the start of the programme




Notes: Earnings measure the individual's monetary earnings in all activities outside the household farm in the last 12 months.
We calculated earnings per month worked from the actual months worked outside the family farm, while total earnings sums earnings over 12 months.
We demeaned outcomes by subtracting the age, strata and regional-specific mean.
We trimmed five per cent highest outliers for presentational purposes, but included them in the test of the equality of distributions (for both monthly and total earnings, distributions are significantly different with and without outliers.)

## Appendix A: Sample design

As with other aspects of this evaluation, which built on an existing randomised controlled trial (RCT) and existing data sources, our sample design and power calculations were directly linked to, and therefore limited by, the earlier studies.
The survey sample for the 2009/2010 RPS follow-up survey included:

1. The 2,515 households targeted for interview during the 2000 baseline surveys in the treatment and control areas and the 2002 baseline survey in the comparison area. This is a representative sample, and as such includes households from all social strata. The study region is composed predominantly of poor households with no specific indigenous affiliation. The enumerator training emphasised the need for including all households and limiting refusals.

2. A 'new' sample of 1,008 households in the treatment and control localities that was added to increase the sample size for the two key age groups. We added 504 households (12 from each of the 42 treatment and control localities) with children born between October 2000 and May 2001, leaving us with a target population of 659 in the treatment and 679 in the control areas for children in this age group. We also added 504 households (12 from each of the 42 treatment and control localities) with young adults born between January and June 1989, leaving us with a target population of 653 in the treatment and 625 in the control areas for children in this age group. Both age groups were over-sampled to maximise the difference in the potential length of exposure to the programme at critical ages between the treatment and control groups. The new sample was randomly selected using census data that was collected in treatment and control localities in May 2000, as well as programme administrative data on all newborns after the census in the treatment and control areas. Therefore, the total number of targeted households is 3,523.

3. All households that have split off, including both local and long-distance migrants, and those that contain the original *titular*, an original panel household member under 22 (in 2010) or a child (under 22 in 2010) of an original household member.

## Appendix B: Survey instruments – overview of questionnaires

### Household questionnaire

The household questionnaire was an expanded version of the original RPS questionnaire. It was a multipurpose questionnaire similar to a World Bank Living Standards Measurement Survey. The topics covered include: (a) household roster which collects characteristics on each of the household members; (b) migration and information to track migrants; (c) household characteristics; (d) household expenditures; (e) health: status, utilisation and expenditure by person; (f) education: includes pre-school education and expenditure by person; (g) agricultural and non-agricultural labour; (h) women's reproductive health and birth history and prenatal care utilisation (up to two births); (i) household consumption; (j) agricultural crop production and consumption; animal production and consumption; (k) labour history; (l) social relations; (m) social programme history.

### Individual questionnaire

The individual questionnaire included a number of achievement and cognitive tests, social-personal questions, questions on relationships within the households and anthropometric measurements. It was given to all people age 21 and younger, though the mother of the child was the respondent for children under the age of 15 for all direct questions that are not part of an achievement or cognitive tests.

This questionnaire included a variety of tests chosen specifically to assess different domains of cognitive functioning – social-personal, language, memory, processing, fine and gross motor skills. It is not well known what areas of cognitive functioning might be closely linked to health, nutrition and poverty. We also gave some consideration to using tests from previous surveys both in Nicaragua and elsewhere to permit comparison of results. These include: (a) the Spanish version of the Peabody Picture Vocabulary Test; (b) Denver Developmental Screening Test; (c) Raven's Coloured Matrices; (d) Digit Span Forward and Backwards; and (e) a number of tests on processing speed, short-term and long-term memory.

The achievement tests included: (a) a maths test that measures problem-solving at various levels of difficulty, similar to the Peabody Individual Achievement Test (Markwardt 1989); (b) grade-level appropriate test on: word identification, spelling, reading fluency and maths fluency. All tests included specific stopping rules and base-level rules to guarantee that children or young adults with disadvantaged backgrounds were not being asked to complete tasks that were too high for their level. The achievement tests also had starting points that reflect the correspondent's educational level.

To provide direct evidence to evaluate our hypotheses, and to mitigate respondent fatigue and ensure high-quality data, we did not give all the achievement and cognitive tests to each age group. Given that we hypothesised that receiving the interventions during early childhood might have had an important impact on cognitive development, and that receiving the educational transfer during primary school years might have primarily enhanced learning, we placed greater emphasis on cognitive tests for children under the age of 15 and on academic achievement for those 15 and older.

We also collected information on the home environment relevant to child development, including observational questions from The Home Observation for Measurement of the Environment.

We measured respondents' social-emotional state using the Strength and Difficulties Questionnaire and the Center for Epidemiologic Studies Depression Scale. Both tests are used widely internationally and are available in Spanish (but were adjusted for Nicaraguan Spanish). By asking these questions directly to the young girls and boys aged 15–21, we captured the opinions of a group of individuals who would typically not be reflected in a household survey.

**Community questionnaire**

The community questionnaire is also based on the RPS questionnaire but was significantly expanded. Topics included in this questionnaire were: (a) history of social programmes; (b) infrastructure; (c) availability and prices of goods; (d) social capital; (e) access and characteristics of education services in the community; (f) access and characteristics of education services in and near the community.

*School facility component:* This questionnaire collected quality information on all the schools and teachers in the survey communities.

*Health facility component:* This questionnaire collected quality information on all public health facilities in or near communities and from the main health centre in the municipality capital.

**Qualitative interview guides**

We developed qualitative field work interview guides for various beneficiaries and others linked to the programme or its objectives. Topics covered included the general context of the community, the RPS programme in general, nutrition, health, education, income sources, migration, other potential programme impacts and other social programmes. We developed guides for community focus groups, former programme *promotoras*, teachers and principals, pre-school instructors or *programa de atencion integral a la nineznicaraguense* officials, and medical personal at health posts or clinics.

## Appendix C: Power calculations

As with other aspects of this evaluation, which built on an existing RCT and existing data sources, our sample design and power calculations were directly linked to, and therefore directly limited by, the earlier studies.

We used power calculations to determine the number of individuals to be oversampled in each age group. These calculations assumed a power of 0.9, an effects size of 0.3 standard deviations on the achievement and cognitive tests, and an intra-cluster correlation of 0.05.

The assumptions underlying the power calculations were based on RPS baseline data for the anthropometric and educational attainment outcomes and data from a similar (neighbouring) population in Nicaragua for cognitive functioning (Macours *et al*. 2012). This latter data set included two of the same cognitive tests – the Spanish version of the Peabody Picture Vocabulary Test and The Denver Developmental Screening Test – for children in the same age groups as those in our study. As suggested by Duflo *et al.* (2008), we were be able to estimate intra-cluster correlations (ICC) based on similar populations within the country, which is preferable to using estimates from other data sources reported in the literature. Using an intra-cluster correlation of 0.05 and a power of 0.9 we are able to detect an effect size of 0.3 standard deviations. The detectable effect size is 0.25 if we assume a power of 0.8. The ICC is on the low side compared to academic test scores taken from schools (Duflo *et al*., 2008). Our ICC likely differs because our cluster is not a school or a class in a school, as is the case for many studies of academic test scores. Rather our cluster is a locality, which in our sample generally contains more than one, and often contains many, different communities. In addition, the type of information we gather on the cognitive tests is often not taught in school. This is consistent with there being greater variation across people living in the same locality.

## Appendix D: Descriptive statistics

This appendix contains selected descriptive statistics referenced in the main report.

Table D1 Frequency of delay for children under seven in RPS population

| | Social | Language | Fine motor | Gross motor |
|---|---|---|---|---|
| Child is in the lowest 25% of international distribution | | | | |
| All | 0.59 | 0.84 | 0.61 | 0.50 |
| Child is in lowest 10% of international distribution | | | | |
| All | 0.40 | 0.63 | 0.42 | 0.35 |
| 6–35 months | 0.24 | 0.50 | 0.30 | 0.39 |
| 36–59 months | 0.40 | 0.61 | 0.55 | 0.39 |
| 60–83 months | 0.63 | 0.84 | 0.48 | 0.24 |

Notes:

Delays are calculated as compared to international standardised scores. The Denver Tests consist of various tasks, all of them age-standardised. Children are categorised in the lowest 25 per cent (10 per cent) if they are in the lowest 25 per cent (10 per cent) for at least one of the tasks.

Source: RPS 2010 follow-up survey. The sample used in this calculation comprises 3,200 children aged 6–83 months, all of whom are children of individuals who were household members originally interviewed as part of the RPS evaluation in 2000 – in Ciudad Darío, Terrabona, Esquipulas, Tuma La Dalia (Matagalpa), Totogalpa and Yalaguina (Madriz) – or in 2002, in Sebaco, San Isidro, San Ramón, Rancho Grande (Matagalpa), Santa Lucia (Boaco) and Palacaguina (Madriz). The tests for early childhood development took place in 2010, in the individual's residence.

We assessed all children's social-personal, language, fine motor and gross motor skills using the four sub-scales of the Denver Developmental Screening Test (Frankenberg and Dodds 1996). As the CLAP, which is widely used in Nicaragua for child monitoring, is a slightly modified version of the Denver, the test is appropriate for the population we study. The Denver can be applied to children covering the entire age range of interest, allowing for comparison between age groups.

Table D2: Balance between early and late treatment (boys aged 9–12 at the programme start)

| | Mean late treatment | Mean early treatment | P-value difference |
|---|---|---|---|
| **Attrited boys** | | | |
| Age at start of transfer | 11.38 | 11.33 | (0.791) |
| Estimated per capita expenditures | 2,750 | 2,469 | (0.182) |
| Household size | 7.404 | 8.198 | (0.211) |
| Number of rooms | 1.615 | 1.429 | (0.355) |
| Animal ownership | 0.115 | 0.148 | (0.547) |
| Distance to primary school (min) | 26.90 | 34.32 | (0.324) |
| Log (land size) | 7.725 | 5.829 | (0.009) |
| Son of household head | 0.788 | 0.776 | (0.915) |
| Father not in household | 0.327 | 0.460 | (0.146) |
| Mother not in household | 0.173 | 0.157 | (0.851) |
| Years of education | 1.365 | 0.978 | (0.286) |
| Enrolled in school | 0.620 | 0.587 | (0.759) |
| Worked in last week | 0.250 | 0.337 | (0.266) |
| Number of observations | 52 | 64 | |
| **Non-attrited boys** | | | |
| Age at start transfer | 11.04 | 10.94 | (0.051) |
| Estimated per capita expenditures | 2341 | 2394 | (0.351) |
| Household size | 8.407 | 8.108 | (0.181) |
| Number of rooms | 1.559 | 1.579 | (0.786) |
| Animal ownership | 0.194 | 0.138 | (0.060) |
| Distance to primary school (min) | 22.39 | 27.443 | (0.321) |
| Log (land size) | 8.128 | 8.130 | (0.995) |
| Son of household head | 0.872 | 0.859 | (0.572) |
| Father not in household | 0.182 | 0.204 | (0.398) |
| Mother not in household | 0.076 | 0.081 | (0.789) |
| Years of education | 1.151 | 1.244 | (0.519) |
| Enrolled in school | 0.763 | 0.752 | (0.826) |
| Worked in last week | 0.251 | 0.244 | (0.837) |
| Number of observations | 499 | 526 | |

Baseline information from RPS census 2000.

## Appendix E: Qualitative study design and methods

*Objective of the preparatory qualitative study*

The objectives of the preparatory qualitative study conducted in 2009 were to understand how the programme operated and what its effects were, from the point of view of the beneficiaries, persons involved in programme execution and others involved directly or indirectly in the programme, such as teachers and health workers. Another crucial objective was to explore why any such effects may have occurred, that is, why and via what mechanisms within the household or community. These findings were then used to refine the research hypotheses about anticipated and unanticipated longer-term impacts of the RPS programme, and in particular the possible mechanisms by which they have come about.

Other objectives of the preparatory qualitative study were to:

- design and pilot additional new questions for the quantitative study
- test cognitive and achievement tests for children and adolescents aged 7–20
- collect information necessary for the closed quantitative survey – for example, names and types of NGOs operating in the areas; crops cultivated; cropping seasons across the different areas.

*Design and methods*

The principal methods used in the study were semi-structured interviews with individuals and in focus groups. Interviewees included: beneficiaries (particularly mothers of children benefitting directly from the programme), voluntary *promotoras* who worked with the programme, teachers, health workers, leaders and others involved in the execution of the programme.

The communities were selected to ensure heterogeneity in terms of their degree of access to educational and health services, economic activities, poverty and agro-climatological characteristics. In each of four mini-regions selected, we visited three communities that were as similar as possible, based on poverty and malnutrition estimates from 2000, to: the original control group, the original intervention group and the comparison group. In addition, we visited four programme municipalities with municipal-level education and health service providers and former programme functionaries.

While we made an effort to capture heterogeneity across community types in the coverage of the qualitative work, it is important to underscore that, given the design of the qualitative study it is not possible in general to make definitive conclusions about the impacts of the programme or the themes examined. This is mainly because the sample is neither representative nor sufficiently large.

## Appendix F: Description of tests

The individual survey includes a number of tests to assess cognition and learning achievement. For young adults aged 15–22, we administered three Spanish language and two maths tests. Specifically, we used one grade level-appropriate test on: word identification, spelling, reading fluency and maths fluency. There was also a second maths test to measure problem-solving at various levels of difficulty, similar to the Peabody Individual Achievement Test (Markwardt 1989). We also administered two tests likely to capture both achievement and cognitive development: the Spanish version of the Peabody Picture Vocabulary Test and a forward and backward digit-span test, in which the respondent is asked to repeat series of numbers read to him. The last test, the Raven Coloured Matrices Test (the 36-item version with sets A, AB and B) likely captures mostly cognition (Raven 1984, 2000), and has been used in many other studies, including in developing countries, for this purpose.[16] We piloted all the tests extensively, and made adjustments for local context as necessary, rephrasing instructions as needed for maximum understanding in the study population.

Very similar tests have been applied in other populations in Latin America, including in cash transfer programme evaluations in Ecuador and Mexico, and a different CCT programme in Nicaragua (Paxson and Schady 2010; Fernald *et al*. 2009; Behrman *et al.* 2009; Macours *et al*. 2012). An important advantage of all the tests is that they provide observed, as opposed to self-reported, measures of learning and cognition, therein substantially reducing concerns about reporting biases.

All tests were conducted in respondents' homes by a specially trained team of test administrators. The results were therefore obtained independently of whether the respondent was in school, avoiding potential selection concerns. During the test administrators' training, great emphasis was placed both on gaining respondents' confidence before starting the tests and on the standardised application of each of the tests. We closely monitored the quality and standardised application of the tests in the field. Given the long survey period, we also organised several re-standardisation trainings, and we organised the data collection and test administration in such a way that test administrators would maintain a balance between the number of children visited in early and late treatment localities. We also balanced visits to early- and late-treatment localities over time to avoid seasonal differences in measurement between the experimental groups. Consistent with these field protocols, results are robust to controls for the identity of the test administrator.

[16] There is a concern that schooling may affect cognitive functioning. The Raven Test does not vary much by schooling in western countries where it has been normed (Raven 2000). Studies that claim that education may affect cognition note that a major limitation of the study is that they are not causal (Ceci 1991). It is highly likely that there is an association between attending school and level of education and cognition in developing countries: who goes to school and how long they go for is likely to be associated with cognitive ability and other socio-economic factors that affect both cognition and academic achievement. Indeed, the present study may be one of the better examples of how an increase in educational attainment of older children does not affect cognitive functioning as measured by the Raven Test.

## Appendix G: Non-experimental methods

Taking advantage of the experiment allows comparison of the early and late treatment groups and provides a differential effect between the two areas. We can also take advantage of a non-experimental comparison group to examine the absolute programme effect of some outcomes. Starting in 2002, data was also collected in the comparison area, before the late treatment group phased into the programme, allowing us to use non-experimental methods to compare the late treatment and comparison group outcomes in 2010. In particular, we use nearest neighbour matching for the first and second neighbour (NN1 and NN5), and use a combination of local smooth (kernel and LLR) and propensity score methods (see Barham *et al.* 2012 for more details on these methods).

We note two important caveats. First the non-experimental estimates cover a different calendar period from the experimental period since the late treatment group received transfers between 2003–2005. It is possible that the differential and absolute impacts may differ depending on other trends in the environment. A possible important difference was the onset of the coffee crisis that reduced overall consumption levels in 2002. Second, there were some modifications to the programme, including reductions in the absolute value of the food security transfers. Arguably, these differences in part offset each other, as they imply that the relative size of the transfers was similar. We therefore estimate consumption impacts in terms of percentage changes. These caveats *a priori* seem less relevant for the impacts on education, as conditionalities, transfer size and implementation of the education component were essentially constant across time.

We present non-experimental groups for two household-level variables – log per capita consumption and food share – and for education and test outcomes presented in Section 5.2 for boys aged 9–12 in 2003 when the late treatment group started to receive the programme. We do not present the non-experimental results on education for the older girls (aged 13 to 14) because the sample size is small as it does not include the oversampled 11-year-olds, making the non-experimental methods difficult to use. In addition, test data is not available for these girls, since we did not collect any for children older than 12 at the start of the programme. Similarly, we do not replicate the results presented in Section 5.1 since the sample sizes are again small.

In Figures G1 and G2 below we present the distributions of the estimated household- and individual-level propensity score models. Observations above the x-axis are from the late treatment group, and those below are from the comparison group. In contrast to what we might have seen if the groups had been randomly allocated, the overlap is imperfect, while substantial, for both models. The late treatment is also skewed to the right while the comparison group is skewed to the left. The various non-experimental estimators address and correct for this difference in the distributions. We restrict the analysis to a common support between the late treatment and comparison groups. We report results from what we refer to as the one per cent trim. In this trimming method, we determine the propensity score of the observation in the late treatment group that is at the first percentile of the late treatment group distribution – in other words, one per cent in from the minimum value but only using the late treatment group – and trim all observations that have a lower propensity score. We do the same at the upper end of the distribution, using the comparison group. We determine the propensity score of the observation in the comparison group that is at the 99th percentile of the comparison group distribution and trim all observations that have a propensity score higher than that value.

Figure G 1: Household-level propensity score distribution

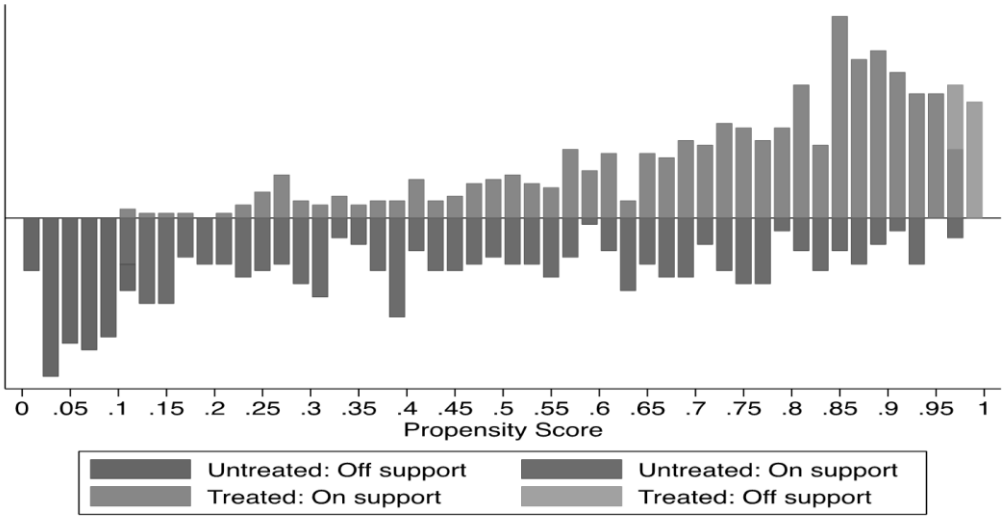
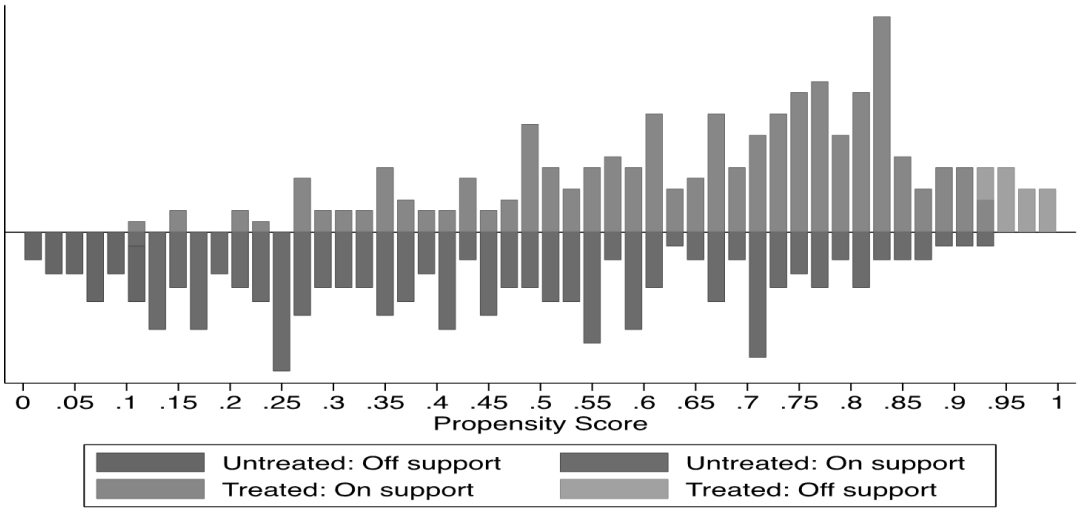



Figure G 2: Individual-level propensity score distribution for boys 9–12 in 2003




The models we used to estimate the propensity scores are shown in Tables G1 and G2 below. There are still some issues of balance and potentially an over-specified model for the number of observations; we are working to improve these scores and hope to report new results in the future. These models do give us a good idea of what the future results may be.

Table G1: Logit results for propensity score matching at household level

| | Late treatment (=1) |
|---|---|
| Log per capita expenditure (Córdobas) | 0.264 |
| | (0.195) |
| *Share of expenditure on:* | |
| Food | −2.941*** |
| | (0.880) |
| Health | −0.775 |
| | (1.619) |
| Education | −5.350** |
| | (2.085) |
| *Share of food expenditures on:* | |
| Staples | 1.254 |
| | (0.833) |
| Animal proteins | 6.008* |
| | (3.487) |
| Animal proteins ^2 | −17.161 |
| | (13.584) |
| Animal proteins ^3 | 17.381 |
| | (15.053) |
| Fruit and vegetables | −0.431 |
| | (1.362) |
| Expenditure on transportation last week (=1) | 0.263 |
| | (0.220) |
| Expenditures on transportation last week (Córdobas) | −0.000** |
| | (0.000) |
| Household head age (log years) | 0.367 |
| | (0.501) |
| Household head male (=1) | −1.077 |
| | (0.673) |
| Household head years of education | 0.104* |
| | (0.058) |
| Household head divorced or widowed (=1) | −0.111 |
| | (0.414) |
| Household head literate (=1) | −0.423 |
| | (0.268) |
| Titular household head (=1) | −0.987 |
| | (0.652) |
| Titular log age * titular log age | −2.972*** |
| | (0.797) |
| Titular log age * titular female=0 | 21.536*** |
| | (5.822) |
| Titular log age * titular female=1 | 21.269*** |
| | (5.829) |
| Titular years of education | 0.142** |
| | (0.057) |
| Titular literate | −0.459* |
| | (0.275) |
| Family size | 0.106* |
| | (0.058) |

| | |
|---|---|
| *Share of household members aged:* | |
| 0–5 and male | −4.912*** |
| | (1.610) |
| 0–5 and female | −4.266*** |
| | (1.613) |
| 6–13 and male | −5.143*** |
| | (1.714) |
| 6–13 and female | −4.814*** |
| | (1.719) |
| 14–17 and male | −2.473 |
| | (1.676) |
| 14–17 and female | −3.022* |
| | (1.691) |
| 18–25 and male | −2.107 |
| | (1.557) |
| 18–25 and female | 0.259 |
| | (1.529) |
| 26–64 and male | −3.528** |
| | (1.598) |
| 26–64 and female | −3.098** |
| | (1.410) |
| 65+ | −2.674 |
| | (1.940) |
| Child aged 6–13 in household (=1) | −0.272 |
| | (0.322) |
| Share of household members age 6–13 enrolled in school | 0.783*** |
| | (0.243) |
| Household has work animals (=1) | −0.414* |
| | (0.244) |
| Household has piped water (=1) | 0.522** |
| | (0.229) |
| Household has well water (=1) | 0.231 |
| | (0.197) |
| Household has a proper kitchen (=1) | 0.063 |
| | (0.220) |
| Household owns house (=10 | −0.393** |
| | (0.193) |
| Household received an electric bill (=1) | −0.580** |
| | (0.280) |
| Household has a radio (=1) | −0.253 |
| | (0.181) |
| Household has an iron (=1) | 1.184*** |
| | (0.448) |
| Household has a blender (=1) | −0.463** |
| | (0.202) |
| Household has a cassette player (=1) | −0.195 |
| | (0.258) |
| Anyone sick in household in past month (=1) | −0.641*** |
| | (0.220) |
| Share of household members working for pay or owner in agriculture | 1.081*** |
| | (0.352) |
| Receive any money from a migrant (=1) | −0.827** |
| | (0.339) |
| Electric bill interacted with iron ownership (=1) | −1.977*** |
| | (0.549) |
| Log 1995 marginality index | 11.949*** |
| | (1.192) |
| Log distance to the municipality capital | −0.013 |
| | (0.154) |
| Constant | −88.929*** |
| | (12.130) |
| Observations | 1,023 |

Notes: Standard errors are clustered at the locality level and in parentheses. ***p < 0.01, ** p < 0.05, *p < 0.10.

Table G2: Logit results for propensity score matching at individual level, boys aged 9–12 in 2003

| | Late treatment (=1) |
|---|---|
| Years of education | −0.184 |
| | (0.114) |
| Days of school missed last month | −0.005 |
| | (0.016) |
| Worked last week (=1) | −0.631** |
| | (0.319) |
| Mother is a household member (=1) | −0.357 |
| | (0.561) |
| Father is a household member (=1) | −0.834 |
| | (0.524) |
| Mother's years of education | 1.501** |
| | (0.706) |
| Log mother's age | 0.508 |
| | (0.813) |
| Mother's years of education * log mother's age | −0.400** |
| | (0.197) |
| Father's years of education | −0.081 |
| | (0.082) |
| Log father's age | −0.404 |
| | (0.831) |
| Household has work animals (=1) | −0.440 |
| | (0.330) |
| Log per capita expenditure (Córdobas) | 0.897*** |
| | (0.251) |
| *Share of expenditure on:* | −2.189* |
| Food | (1.199) |
| | −4.468** |
| Health | (2.038) |
| | −6.895** |
| Education | (3.080) |
| | 1.641** |
| Household head age (log years) | (0.707) |
| | −0.879 |
| Household head male (=1) | (0.944) |
| | 0.099 |
| Household head years of education | (0.110) |
| | −0.603 |
| Household head is literate (=1) | (0.368) |
| | −0.318 |
| Titular household head (=1) | (0.902) |
| | 0.067 |
| Titular literate | (0.351) |
| | 0.068 |
| Family size | (0.058) |
| | 0.811** |
| Household has piped water (=1) | (0.328) |
| | 0.472* |
| Household has well water (=1) | (0.281) |

| | |
|---|---|
| | 0.136 |
| Household has a proper kitchen (=1) | (0.311) |
| | −0.274 |
| Household owns house (=10 | (0.275) |
| | −0.529 |
| Household received an electric bill (=1) | (0.452) |
| | −0.000 |
| Household has a radio (=1) | (0.257) |
| | 1.716*** |
| Household has an iron (=1) | (0.615) |
| | −0.507 |
| Household has a blender (=1) | (0.316) |
| | −0.487 |
| Household has a cassette player (=1) | (0.412) |
| | −0.782 |
| Receive any money from a migrant (=1) | (0.497) |
| | 0.162 |
| Anyone sick in household in past month (=1) | (0.329) |
| | −2.142*** |
| Electric bill interacted with iron ownership (=1) | (0.798) |
| | 0.086*** |
| 1995 marginality index | (0.020) |
| | −0.131 |
| Log distance to the municipality capital | (0.230) |
| | 1.861*** |
| Share of household members working for pay or owner in agriculture | (0.533) |
| | −18.672*** |
| Constant | (4.039) |
| | −0.184 |
| Observations | 440 |
| R-squared | |

Standard errors in parentheses ***p < 0.01, **p < 0.05, *p < 0.1

## References


Adato, M and Roopnaraine, T, 2004. *A social analysis of the Red de Protección Social.* Report submitted to the Red de Protección Social. Washington, DC.: International Food Policy Research Institute. Photocopy.

Almond, D and Currie, J, 2010. *Human capital development before age five.* NBER working paper No. w15827.

Attanasio, O, Gómez, L, Heredia, P and Vera-Herández, M, 2005. *The short-term impact of conditional cash subsidy on children's health and nutrition in Colombia*. Unpublished manuscript. London: University College.

Baird, S, Ferreira, F, Özler, B and Woolcock, M, 2012. *Relative effectiveness and cost-effectiveness of conditional and unconditional cash transfers for schooling outcomes in developing countries: a systematic review. PROTOCOL.* The Campbell Collaboration. Available at: <www.campbellcollaboration.org/lib/download/2275>

Barham, T, 2012. The effects of child health interventions on later life cognitive functioning: evidence from Matlab in Bangladesh. *American Economic Journal: Applied Economics*, 4(1), pp.245–73.

Barham, T, 2011. Providing a healthier start to life: the impact of conditional cash transfer programmes on infant mortality. *Journal of Development Economics*, 94(1), pp.74–85.

Barham, T, Di Maro, V, Macours, K and Maluccio, JA, 2012. *Assessing non-experimental impact estimates using randomized phase-in*. Mimeo.

Barham, T, Macours, K and Maluccio, JA, 2013a. Boys' cognitive skill formation and physical growth: long-term experimental evidence on critical ages for early childhood interventions, *American Economic: Papers and Proceedings*, 103(3), pp.467–71 (and IADB working paper).

Barham, T, Macours, K and Maluccio, JA, 2013b. *More schooling and more learning? effects of a 3-year conditional cash transfer programme in Nicaragua after 10 years*. IADB working paper.

Barham, T and Maluccio, JA, 2009. Eradicating disease: the effect of conditional cash transfers on vaccination coverage in Nicaragua. *Journal of Health Economics,* 38(3), pp. 611–21.

Bautista-Arredondo, S, Bertozzi, S, Gertler, P and Martinez, S, 2006. *How healthcare supply constraints limit the effectiveness of conditional cash transfer programmes: evidence from Mexico's Oportunidades program.* Unpublished manuscript. Berkeley: University of California.

Behrman, J and Hoddinott, J, 2005. Programme evaluation with unobserved heterogeneity and selective implementation: the Mexican PROGRESA impact on child nutrition. *Oxford Bulletin of Economics and Statistics*, 67(4), pp.547–69.

Behrman, JR, Parker, SW and Todd, PE, 2009. Medium-term impacts of the *Oportunidades* conditional cash transfer programme on rural youth in Mexico. In: S Klasen and F Nowak-Lehmann, eds. 2009. *Poverty, inequality, and policy in Latin America.* Cambridge: MIT Press, pp.219–70.

Behrman, JR, Parker, SW and Todd, PE, 2011. Do conditional cash transfers for schooling generate lasting benefits? A five-year follow-up of PROGRESA/*Oportunidades*. *Journal of Human Resources*, 46(1), pp.93–122.

Bryan, G, Chowdhury, S and Mobarak, M, 2013. *Escaping famine through seasonal migration.* London School of Economics. Mimeo.

Card, D. 1999. The causal effect of education on earnings. In: O. Ashenfelter and D. Card, eds. *The handbook of labor economics*, Vol. 3B, New York: North-Holland.

Card, D, 2001. Estimating the return to schooling: progress on some persistent econometric problems. *Econometrica*, 69(5), pp.1,127–160.

Case A, Fertig, A and Paxson, C, 2005. The lasting impact of childhood health and circumstance. *Journal of Health Economics*, 24(2), pp.365–89.

Ceci, SJ, 1991. How much does schooling influence general intelligence and its cognitive components? a reassessment of the evidence. *Developmental Psychology*, 27(5), pp.703–22.

Cunha, F, Heckman, J, Lochner, L and Masterov, D, 2006. Interpreting the evidence on life cycle skill formation. In: E Hanushek and F Welch, eds. *The handbook of the economics of education*, Vol 1, pp.697–812. Amsterdam: North-Holland Press.

De Janvry, A, Finan, F, Sadoulet, E and Vakis, R, 2006. Can conditional cash transfers serve as safety nets in keeping children at school and from working when exposed to shocks? *Journal of Development Economics*, 79(2), pp.349–73.

Duflo, E, Glennerster, R and Kremer, M, 2008. Using randomization in development economics research: a toolkit. In: TP Schulz and JA Strauss, eds. *Handbook of Development Economics* 4, pp.3895–962. Amsterdam, The Netherlands: Elsevier.

Duflo, E, 2001. Schooling and labour market consequences of school construction in Indonesia: evidence from an unusual policy experiment. *American Economic Review,* 91, pp.795–813.

Dunn, LM, Lugo, DE, Padilla, ER and Dunn, L, 1986. *Test de Vocabulario en Imágenes Peabody*. Circle Pines, Minnesota, US: American Guidance Service.

Edmonds E and Schady, N, 2012. Poverty alleviation and child labour. *American Economic Journal: Economic Policy,* 4(4), pp.100–24.

Eriksson, JG, Kajantie, E, Osmond, C, Thornburg, K and Barker, DJP, 2010. Boys live dangerously in the womb. *American Journal of Human Biology,* 22, pp.330–35.

Fernald, L, Gertler, P and Neufeld, L, 2009. 10-year effect of *Oportunidades*, Mexico's conditional cash transfer programme, on child growth, cognition, language, and behaviour: a longitudinal follow-up study. *The Lancet,* 371, pp.828–37.

Filmer, D, Hasan, A and Pritchett, L, 2006. *A millennium learning goal: measuring real progress in education.* Center for Global Development working paper 97.

Filmer, D, and Schady, N, 2014. The medium-term effects of scholarships in a low-income country. *Journal of Human Resources* (in press).

Fiszbein, A and Schady, N, 2009. *Conditional cash transfers: reducing present and future poverty.* World Bank policy research report. Washington DC: The World Bank.

Frankenberg, WK and Dodds, JB. 1996. The Denver Developmental Screening Test. Denver, CO: University of Colorado Medical Center.

Gaarder, MM, Glassman, A and Todd, JE, 2010. Conditional cash transfers and health: unpacking the causal chain. *Journal of Development Effectiveness*, 2:1, pp.6–50.

Garces, E., Thomas, D and Currie, J, 2002. Longer-term effects of head start. *American Economic Review*, 92(4), pp.999–1,012.

Gertler, P. 2004. Do conditional cash transfers improve child health? evidence from PROGRESA's control randomized experiment. *American Economic Review*, 94(2), pp. 336–41.

Gertler, P, Martinez, S and Rubio-Codina, M, 2012. Investing cash transfers to raise long term living standards. *American Economic Journal: Applied Economics,* Vol 4(1), pp. 164–92.

Glewwe, P, 2002. Schools and skills in developing countries: education policies and socioeconomic outcomes. *Journal of Economic Literature,* Vol XL, pp.436–82.

Glewwe, P and Kremer, M, 2006. Schools, teachers, and education outcomes in developing countries. In: E Hanushek and F Welch, eds. *Handbook of the economics of education*, Vol 2, pp.945–1,017. Amsterdam: North-Holland Press.

Glewwe, P, Holla, A and Kremer, M, 2008. *Teacher incentives in the developing world.* Unpublished manuscript. Cambridge MA: Harvard University.

Grantham-McGregor, S, Walker, S, Chang, S and Powell, C, 1997. Effects of early childhood supplementation with and without stimulation on later development in stunted Jamaican children. *American Journal of Clinical Nutrition*, 66, pp.247–53.

Griliches, Z, 1977. Estimating the returns to schooling: some econometric problems. *Econometrica*, 45(1), pp.1–22.

Heckman, J, 2006. Skill formation and the economics of investing in disadvantaged children. *Science*, 31, pp.1,900–902.

Heckman, J and Smith, JA, 1995. Assessing the case for social experiments. *Journal of Economic Perspectives,* 9(2), pp.85–110.

Hoddinott, J, Maluccio, JA, Behrman, JR, Flores, R and Martorell, R, 2008. Effect of a nutrition intervention during early childhood on economic productivity in Guatemalan adults. *The Lancet*, 371 (9610). pp.411–16.

IEG (Independent Evaluation Group). 2011. *Evidence and lessons learned from impact evaluations on social safety nets.* Washington, DC: The World Bank.

Kling, J, Liebman, J and Katz, L, 2007. Experimental analysis of neighborhood effects. *Econometrica* 75(1), pp.83–119.

Levy, S and Schady, N, 2013. Latin America's social policy challenge: education, social insurance, redistribution. *Journal of Economic Perspectives,* 27(2), pp.193–218.

Macours, K and Vakis, R, 2010. Seasonal migration and early childhood development. *World Development*, 38(6), pp.857–69.

Macours, K, Schady, N and Vakis, R, 2012. Cash transfers, behavioural changes, and cognitive development in early childhood: evidence from a randomised experiment. *American Economic Journal: Applied Economics*, 4(2), pp.247–73.

Maluccio, JA, 2009. Household targeting in practice: The Nicaraguan *Red de Protección Social*. *Journal of International Development*, 21(1), pp.1–23.

Maluccio, JA, 2010. The impact of conditional cash transfers on consumption and investment in Nicaragua. *Journal of Development Studies*, 46(1), pp.14–38.

Maluccio, JA and Flores, R, 2005. *Impact evaluation of a conditional cash transfer program: The Nicaraguan Red de Protección Social.* Research report 141. Washington DC: International Food Policy Research Institute.

Maluccio, JA, Hoddinott, J, Behrman, JR, Martorell, R, Quisumbing, A and Stein, AD, 2009. The impact of experimental nutritional interventions on education into adulthood in rural Guatemala. *Economic Journal*, 119 (April), pp.734–63.

Maluccio, JA, Murphy, A and Regalia, F, 2010. Does supply matter? initial school conditions and the effectiveness of conditional cash transfers for grade progression in Nicaragua. *Journal of Development Effectiveness,* 2(1), pp.87–116.

Markwardt, FC, 1989. *Peabody Individual Achievement Test – revised manual*. Circle Pines, Minnesota, US: American Guidance Service.

Mosley, W and Chen, L, 1984. An analytical framework for the study of child survival in developing countries. *Population and Development Review*, 10, pp.25–45.

National Research Council. 1986. *Population growth and economic development: policy questions*. Washington, DC: National Academy Press.

Oliver, R, 1999. Fertility and women's schooling in Ghana. In: P Glewwe, ed. *The economics of school quality investments in developing countries*. New York: St. Martin's, Macmillan, pp.327–44.

Parker, S, Rubalcava, L and Teruel, G, 2008. Evaluation of conditional schooling and health programmes. In: TP Schultz and J Strauss, eds. *Handbook of development economics*, Vol. 4. Amsterdam: North-Holland Press, pp.3,964–980.

Paxson, C, and Schady, N, 2010. Does money matter? the effects of cash transfers on child health and development in rural Ecuador. *Economic Development and Cultural Change,* 59(1), pp.187–230.

Pollitt, E, Gorman, KS, Engle, P, Martorell, R and Rivera, JA, 1993. Early supplementary feeding and cognition: effects over two decades. *Monographs of the Society for Research in Child Development,* Serial No 235, 58 (7).

Raven, JC, Court, JH and Raven, J, 1984. *Manual for Raven's progressive matrices and vocabulary scales, section 2: coloured progressive matrices.* London: HK Lewis.

Raven, J, 2000. The Raven's progressive matrices: change and stability over culture and time. *Cognitive Psychology*, 41, pp.1–48.

Regalia, F, and Castro, L, 2007. *Performance-based incentives for health: demand and supply-side incentives in the Nicaraguan Red de Protección Social*. Washington, DC: Center for Global Development working paper No 119.

Rivera, J, Sotres-Alvarez, D, Habicht, J-P, Shamah, T and Villalpando, S, 2004. Impact of the Mexican programme for education, health, and nutrition (PROGRESA) on rates of growth and anemia in infants and young children - a randomized effectiveness study. *Journal of the American Medical Association*, 291(21), pp.2,563–570.

Schady, N and Araujo, MC, 2008. Cash transfers, conditions and school enrollment in Ecuador. *Economía*, 8(2), pp.43–70.

Schultz, TP, 2004. School subsidies for the poor: evaluating the Mexican PROGRESA poverty program. *Journal of Development Economics*, 74, pp.119–250.

Stillman, S, Gibson, J and McKenzie, D, 2012. *Miserable migrants? Natural experiment evidence on international migration and objective and subjective well-being*. IZA discussion papers 6871. Bonn: Institute for the Study of Labor (IZA).

Strauss, J and Thomas, D 1998. Health, nutrition, and economic development. *Journal of Economic Literature*, 36(2), pp.766–817.

Strauss, J and Thomas, D, 2008. Health over the life course. In: T.P. Schultz and J. Strauss eds. *Handbook of development economics*, Vol 4. New York, NY: North-Holland Press, pp. 3,375–474.

Thomas, D, 1999. Fertility, education and resources in South Africa. In: C. Bledsoe *et al.* eds. *Critical perspectives on schooling and fertility in the developing world.* Washington, DC: National Academy Press.

Walker, S, Wachs, T, Gardner, J, Lozoff, B, Wasserman, GA, Pollitt, E, Carter, JA and the International Child Development Steering Group, 2007. Child development in developing countries 2: child development: risk factors for adverse outcomes in developing countries. *The Lancet* 369 (9556), pp.145–57.

Wicherts, JM, Donlan, CV, Carlson, JS and van der Maas, HLJ, 2010. Raven's Test performance of Sub-Saharan Africans: average performance, psychometric properties, and the Flynn Effect. *Learning and Individual Differences*, 20, pp.135–51.

World Bank, 2013. *World Development Indicators* 2013. Accessed 23 November 2013.

## Publications in the 3ie Impact Evaluation Report Series

The following reports are available from http://www.3ieimpact.org/evidence-hub/impact-evaluation-repository

*The promise of preschool in Africa: A randomised impact evaluation of early childhood development in rural Mozambique*, *3ie Impact Evaluation Report 1*. Martinez, S, Naudeau, S and Pereira, V (2012)

*A rapid assessment randomised-controlled trial of improved cookstoves in rural Ghana*, *3ie Impact Evaluation Report 2*. Burwen, J and Levine, DI (2012)

*The GoBifo project evaluation report: Assessing the impacts of community-driven development in Sierra Leone*, *3ie Impact Evaluation Report 3*. Casey, K, Glennerster, R and Miguel, E (2013)

*Does marginal cost pricing of electricity affect groundwater pumping behaviour of farmers? Evidence from India, 3ie Impact Evaluation Report 4*. Meenakshi, JV, Banerji, A, Mukherji, A and Gupta, A (2013)

*Impact evaluation of the non-contributory social pension programme 70 y más in Mexico*, *3ie Impact Evaluation Report 5.* Rodríguez, A, Espinoza, B, Tamayo, K, Pereda, P, Góngora, V, Tagliaferro, G and Solís, M (2014)

*The impact of daycare on maternal labour supply and child development in Mexico*, *3ie Impact Evaluation Report 6*. Angeles, G, Gadsden, P, Galiani, S, Gertler, P, Herrera, A, Kariger, P and Seira, E (2014)

*Social and economic impacts of Tuungane: final report on the effects of a community-driven reconstruction programme in the Democratic Republic of Congo*, *3ie Impact Evaluation Report 7*. Humphreys, M, Sanchez de la Sierra, R and van der Windt, P (2013)

*Paying for performance in China's battle against anaemia*, *3ie Impact Evaluation Report 8*. Zhang, L, Rozelle, S and Shi, Y (2013)

*No margin, no mission? Evaluating the role of incentives in the distribution of public goods in Zambia*, *3ie Impact Evaluation Report 9*. Ashraf, N, Bandiera, O and Jack, K (2013)

*Truth-telling by third-party audits and the response of polluting firms: Experimental evidence from India, 3ie Impact Evaluation Report 10.* Duflo, E, Greenstone, M, Pande, R and Ryan, N (2013)

*An impact evaluation of information disclosure on elected representatives' performance: evidence from rural and urban India, 3ie Impact Evaluation Report 11.* Banerjee, A, Duflo, E, Imbert, C, Pande, R, Walton, M and Mahapatra, B (2014)

*Targeting the poor: evidence from a field experiment in Indonesia, 3ie Impact Evaluation Report 12.* Atlas, V, Banerjee, A, Hanna, R, Olken, B, Wai-poi, M and Purnamasari, R (2014)

*Scaling up male circumcision service provision: results from a randomised evaluation in Malawi, 3ie Impact Evaluation Report 13.* Thornton, R, Chinkhumba, J, Godlonton, S and Pierotti, R (2014)

*Providing collateral and improving product market access for smallholder farmers: a randomised evaluation of inventory credit in Sierra Leone, 3ie Impact Evaluation Report 14.* Casaburi, L, Glennerster, R, Suri, T and Kamara, S (2014)

*A youth wage subsidy experiment for South Africa, 3ie Impact Evaluation Report 15.* Levinsohn, J, Rankin, N, Roberts, G and Schöer, V (2014)

*The impact of mother literacy and participation programmes on child learning: evidence from a randomised evaluation in India, 3ie Impact Evaluation Report 16.* Banerji, R, Berry, J and Shortland, M (2014)

*Assessing long-term impacts of conditional cash transfers on children and young adults in rural Nicaragua*, *3ie Impact Evaluation Report 17*. Barham, T, Macours, K, Maluccio, JA, Regalia, F, Aguilera, V and Moncada, ME (2014)

*Impact of malaria control and enhanced literacy instruction on educational outcomes among school children in Kenya: a multi-sectoral, prospective, randomised evaluation, 3ie Impact Evaluation Report 18*. Brooker, S and Halliday, K (2014)

*A randomised evaluation of the effects of an agricultural insurance programme on rural households' behaviour: evidence from China, 3ie Impact Evaluation Report 19.* Cai, J, de Janvry, A and Sadoulet, E (2014)

*Environmental and socioeconomic impacts of Mexico's payments for ecosystem services programme*, *3ie Impact Evaluation Report 20*. Alix-Garcia, J, Aronson, G, Radeloff, V, Ramirez-Reyes, C, Shapiro, E, Sims, K and Yañez-Pagans, P (2014)

*Shelter from the storm: upgrading housing infrastructure in Latin American slums*, *3ie Impact Evaluation Report 21*. Galiani, S, Gertler, P, Cooper, R, Martinez, S, Ross, A and Undurraga, R (2014)

While there is evidence on the positive short-term impacts of conditional cash transfer (CCT) programmes on education, health and poverty reduction, little is known about their long-term effects. This research studies the impacts of *Red de Protección Social* (Social protection network or RPS), a CCT programme targeting rural households in impoverished regions in Nicaragua from 2000 to 2005. RPS involved two household-level transfers: for health and nutrition and for educating school-age children. The results show that CCTs can lead to gains in human capital accumulation that last beyond the life of the programme, suggesting they have potential to lead to long-term poverty reduction.

**Impact Evaluation Series**

International Initiative for Impact Evaluation
c/o Global Development Network
2nd Floor, West Wing, ISID Complex
Plot No. 4, Vasant Kunj Institutional Area
New Delhi – 110070
India

3ie@3ieimpact.org
Tel: +91 11 4323 9494

**www.3ieimpact.org**