Essays on Development Economics

Julieta M. Trias

June 9, 2013

A thesis submitted for the Degree of Doctor of Philosophy to the University College London, the University of London.

Declaration

I, Julieta M. Trias, confirm that:

• the work presented in this thesis is my own and it has not been presented to any other university or institution for a degree,

• where information has been derived from other sources, I confirm that this has been indicated in the thesis.

• chapter one is a sole authored paper.

• chapter two is a sole authored paper.

• chapter three is based on conjoint work with Orazio Attanasio (University College London) and Marcos Vera Hernandez (University College London).

Julieta Trias

Signature

June 9, 2013

Date

Acknowledgements

I would like to thank my supervisors Marcos Vera Hernandez and Orazio Attanasio for being a constant source of support, encouragement, and inspiration during my entire Ph.D.

Special thanks to Samuel Berlinski who was an important adviser during the initial period.

To Leonardo Gasparini who has been my inspiration in my academic and professional life and who encourage me to pursuit doctoral studies.

To my friends and colleges at UCL for invaluable comments to my research and for their company and understanding at every stage of the Ph.D. Particularly, to Laura Jaitman, Christian Krestel, Renata Narita, Abhishek Chakravarty, Brendon McConnel, Gunilla Petterson, Elaine Kelly, Italo Garcia, Cristina Santos and Marieke Schnabel. Special thanks to Paula Giovagnoli who encourage me to come to UCL and gave me all her support during all my PhD.

I would like to thank my husband Alejandro and my parents for helping me going through the most difficult moments, for their company, support and patience over all these years.

Abstract

Chapter 1 investigates the impact of weather-related income shocks on infant mortality in rural Ecuador. I find that favorable weather conditions during the growing season have a negative effect on infant survival rates when the harvest takes place during the first and third trimester of pregnancy, and the first trimester following birth. My results suggest that the negative effects of an increased maternal labour supply, following a positive agricultural productivity shock, during pregnancy and the first trimester after birth outstrip the positive effects resulting from the consequent higher income when considering year-to-year weather fluctuations. I also find that favorable weather during the growing season reduces -via maternal time- prenatal care, skilled assistance at birth, and breastfeeding duration and frequency. Chapter 2 explores the presence of spillover effects on schooling outcomes from the Colombian welfare program, "Familias en Acción", on ineligible households in rural areas. The program provides cash subsidies to poor families conditional on children school attendance. I find that ineligible children — those living in a household that has not been classified as poor—residing in targeted areas are more likely to stay in the school during the transition period between primary and secondary school. My results suggest that *peer effects* might play an important role in schooling decisions as the increased grade completion rate of the peer group increases the individual completion. Chapter 3 uses a randomized experiment to examine the causal effect of improving housing conditions on child health, and adult mental health. We find that replacing floors, upgrading toilets, kitchen, and play areas has no impact on child health but these results are subject to a high level of non-selective attrition on children. We also find that the program improves caregiver's mental health as measured by the CES-D depression score.

Contents

List of Tables

List of Figures

Introduction

This thesis consists of three chapters, two of them focus on early childhood conditions and child health, and the third one explore spillover effects of a program aiming to increase human capital accumulation providing a grant to poor families conditional on school attendance and health check-ups.

Chapter 1 investigates the effects of weather-related transitory income fluctuations during the prenatal and postnatal period on infant mortality in rural Ecuador. The body of literature discussing the effect of income shocks on infant mortality has reported mixed evidence (Ferreira & Schady, 2009). Several transmission mechanisms have been proposed to explain this phenomenon, including reduction in the opportunity cost of woman's time inducing mothers to substitute time away from labour market towards health-preserving activities such as prenatal care, decrease of pollutant emissions, and reduction in the consumption of harmful goods such as smoking and alcohol. In this chapter, I exploit year-to-year variation in rainfall and temperature across time and space to, first, estimate the net effect (positive or negative) on infant mortality of weather-related agricultural shocks taking place during the fetal and postnatal period and; second, to investigate the likely mechanisms behind the measured net effect.

In rural areas where households depend heavily on agriculture, weather fluctuations are a key factor driving the evolution of production and employment. Using monthly data from 340 weather stations across the country from 1995-2009 I construct a proxy for agricultural productivity shocks –crops' ideal temperatures and deviations from average rainfalls during the growing season. However, weather can have extensive effects, affecting not only agricultural production but the transmission of vector-borne diseases. After controlling for malaria-zone, I find that unfavorable (favorable) weather conditions during the growing season have a positive (negative) effect on infant survival when the harvest takes place during the first and third trimester of pregnancy, and the first trimester following birth. Crucial determinants of child health are inexpensive but require large amounts of time (traveling to distant facilities for free prenatal care and primary health services or breastfeeding). My results suggest that the negative effects of an increased maternal labour supply, following a positive agricultural productivity shock, during pregnancy and the first trimester after birth outstrip the positive effects resulting from the consequent higher income when considering year-to-year weather fluctuations. I find that unfavorable (favorable) weather during the growing season increases (reduces) -via maternal time - prenatal care, skilled assistance at birth, and breastfeeding duration and frequency. These effects are consistent with the argument of value of time in the production of child health.

Chapter 2 explores the presence of spillover effects on schooling outcomes from the Colombian welfare program, “Familias en Acción”, on ineligible households in rural areas. The program provides cash subsidies to poor families conditional on children school attendance. The identification strategy exploits the fact that exposure to the program varies by locality of residence and date of birth. Using a 10% census sample from 2005 I find positive spillover effects in areas where the program was available. Ineligible children — those living in a household that has not been classified as poor—residing in targeted areas are more likely to stay in the school during the transition period between primary and secondary school. My estimates report an increment by 9 percentage points in primary school completion rate among non-poor children and 7-10 percentage points for grade completion in the initial years of secondary school. Using the partial population experiment setting created by the intervention I extend the analysis to assess the response of the ineligible children to the introduction of the program in their peer group —children from the same birth cohort living in the same municipality. The results suggest that behavioral social interactions (*peer effect*) might play an important role in schooling decisions where the increased grade completion rate of the peer group increases the individual completion. This finding implies that policies aimed at encouraging enrollment can produce large social multiplier effects.

In chapter 3, we use a randomized experiment to examine the causal effect of improving housing conditions on child health and adult mental health. Housing conditions affects health in many different ways. Deficient housing conditions on aspects such as water, sanitation, and safe food preparation and storage, can be a critical factor behind the rapid spread of communicable and food-borne diseases and consequently the accumulation of human capital. Moreover, housing and neighborhoods can be also considered as a psychosocial environment that can affect mental health and life satisfaction. We use a randomized experiment to examine the causal effect of improving housing conditions on child health and adult mental health. The experiment was carried-out in the city of Cartagena, Colombia, and the intervention consisted in upgrading floors, walls, and toilets. The program was specifically targeted to houses accommodating a state-funded community nursery called *“Hogares Comunitarios”*. We find that replacing floors, upgrading toilets, kitchen, and play areas has no impact on child health but these results are subject to a high level of non-selective attrition on children. We also find that the program improves caregiver’s mental health as measured by the CES-D depression score. Size effects are moderate on average, at around 0.3 to 0.4 standard deviations, with stronger effects for those between the 30 and 70 percentile of the depression score distribution.

1 The positive impact of negative shocks: Infant mortality and agricultural productivity shocks

Abstract

In this paper I investigate the impact of weather-related income shocks on infant mortality in rural Ecuador. Using monthly data from 340 weather stations across the country from 1995-2009, I construct a proxy for agricultural productivity shocks – crops' ideal temperatures and deviations from average rainfalls during the growing season. Based on administrative records, I compute monthly infant mortality rates at the parish level. Controlling for vector borne diseases and agricultural prices, I find that unfavorable (favorable) weather conditions during the growing season have a positive (negative) effect on infant survival when the harvest takes place during the first and third trimester of pregnancy, and the first trimester following birth. Crucial determinants of child health are inexpensive but require large amounts of time (traveling to distant facilities for free prenatal care and primary health services or breastfeeding). My results suggest that the negative effects of an increased maternal labour supply, following a positive agricultural productivity shock, during pregnancy and the first trimester after birth outstrip the positive effects resulting from the consequent higher income when considering year-to-year weather fluctuations. I find that unfavorable (favorable) weather during the growing season increases (reduces) -via maternal time - prenatal care, skilled assistance at birth, and breastfeeding duration and frequency. These effects are consistent with the argument of value of time in the production of child health.

1.1 Introduction

What is the effect of transitory income shocks on infant mortality? Evidence for developed countries suggests that mortality decreases when the economy temporarily weakens (see Ruhm 2004 for a review, Rhum 2000, and Dehejia & Lleras-Muney 2004). Several transmission mechanisms have been proposed to explain this phenomenon, including: reduction in the opportunity cost of women's time inducing mothers to substitute time away from labour market towards health-preserving activities such as prenatal care (Dehejia & Lleras-Muney 2004); decrease of pollutant emissions (Chay & Greenstone, 2003) and; reduction in the consumption of harmful goods such as smoking and alcohol (Dehejia & Lleras-Muney 2004). However, evidence for developing countries on the effects of income shocks on infant mortality is mixed (for a review of such evidence see Ferreira & Schady, 2009). Pritchett & Summers (1996), Cutler *et al.* (2002), Bhalotra (2010) and Baird et al. (2011) report a negative impact from a recession while other researchers find positive associations (Miller and Urdinola, 2010).

Understanding the effect of transitory income fluctuations on infant health is crucial to understand human capital accumulation. There is growing evidence that early life conditions can have persistent and profound impacts on later life. Almond and Currie (2011) present a detailed review on the causal relationship between shocks in early childhood and future outcomes based on shocks affecting maternal and fetal health (nutrition and infections), economic shocks, and pollution.

This paper investigates the effects of transitory income fluctuation during the prenatal and postnatal period on infant mortality in rural Ecuador. I exploit year to year variation in rainfall and temperature across time and space. In rural areas where households depend heavily on agriculture, weather fluctuations are a key factor driving the evolution of production and employment. For children born in rural areas, rainfall and crops' ideal temperature variation across time and space should generate the corresponding variation in agricultural production. The objective of this paper is two-fold: first, it estimates the net effect (positive or negative) on infant mortality of weather-related agricultural shocks taking place during the fetal and postnatal period and; second, it investigates the likely mechanisms behind the measured net effect.

Weather during the planting and growing season determines agricultural production.[1] Households are likely to be affected at the time of the harvest, when employment and income are more affected with the high or low agricultural production. Weather fluctuations during the growing season induce both income and substitution effects on rural households. Unfavorable weather generates a negative income effect that works towards a deterioration of outcomes (via medical inputs and nutrition) - especially if households are credit constrained. The substitution effect works in the opposite direction resulting in improved health status because maternal labour supply declines (and time for health-improving activities increases). The latter can play an important role because crucial determinants of child health are inexpensive but require large amounts of time like traveling to distant facilities for free prenatal care and primary health services or breastfeeding.

This investigation has considerable data requirements. I use birth and death certificates to measure infant mortality by month-year of birth and parish of birth. I combine data from 340 weather station across Ecuador to create proxy variables of weather related productivity shocks - crop specific ideal temperature and rainfall measures during

[1] In this paper, I consider planting and growing season as a unique season and I will refer to it as growing season.

the planting and growing season, and malaria-ideal temperatures and rainfall. [2] Based on a malaria map, I identify parishes with and without malaria. To explore the mechanisms involved and health investments, I use the Living Conditions Surveys (LCS 1998, 1999 and 2004), Reproductive and Health Surveys (RHS 1999 and 2004) and birth certificates (1996-2009).

In this paper I consider that agricultural shocks will affect fetal and infant health at the time of the harvest when employment is more affected by the level of agricultural production. [3] I find that unfavorable weather during the growing-season reduces infant mortality for birth cohorts that were exposed to the agricultural shock during the first and third trimester of the *in utero* period and during the first trimester of life. This implies that the substitution effect seems to be stronger than income effect. Negative shocks during the pregnancy decrease maternal work, prenatal care and increase skilled attendance at birth while shocks after birth reduce breastfeeding duration and frequency. After the third month after birth the income effect seems to outweigh the substitution effect. Estimations take into account that transmission of vector-borne diseases are influenced by environmental and climatic factors and they can affect fetal and infant health. I include a set of variables to control for malaria's ideal weather and other effects of weather on health. Estimations also include controls for external agricultural price shocks.

This paper contributes to the literature on weather and death (Burgess *et al.* (2011) and Kudamatsu (2012)), and early life conditions and human capital accumulation (Banerjee *et al.* (2009), Maccini & Yang (2009)) providing empirical evidence on differential effects of prenatal and postnatal exposure to weather-related income shocks. Exposure to agricultural shocks during the prenatal and neonatal period produces larger effects on infant mortality than later exposure. I show that the main mechanism seems to operate through maternal labour supply -via maternal time. This paper also contributes to the literature on primacy of time in the production of child quality (Mayer (1997); Blau (1999), Price (2008)).

The rest of the paper is organized as follows. Section 1.2 discusses the link between weather and infant mortality. Section 1.3 presents the identification strategy and section 1.4 describes the data used and basic statistics. The results and the mechanisms involved are summarized in section 1.5 and 1.6, and the robustness checks are pre-

[2] It is worth noting that each agricultural, weather, and health variable consider different periods of time. For instance, if the harvest occurs during the first trimester of the pregnancy, agricultural weather related variables will consider the weather before the pregnancy while health weather related variables will consider the weather during the pregnancy.

[3] Crops yield is affected by weather during the growing season. Then unfavorable weather will reduce production at the time of the harvest.

sented in section 1.7. Finally, section 1.8 summarizes and discusses the implications of the results.

1.2 Background: Link between weather, agriculture and infant health

There are several channels through which weather can affect infant mortality. I classify them into two categories: agriculture-related and health-related. In the first channel, weather affects crops productivity. Changes in crops productivity affect wages and the opportunity cost of health preserving activities. Also, mitigating actions can take place and those actions might be harmful for human health (pesticides). In the second channel, weather can have direct impact on human health as well as an indirect one through changes in environmental diseases. This section describes in more detail the possible pathways through which weather can affect infant health.

1.2.1 Weather and Agriculture

Effect on crops productivity There are two mechanisms through which precipitations and temperature can affect crops. The direct channel is mediated by the inputs in the production function (temperature and moisture). *Temperature* is one of the major environmental factors affecting the growth, development and yields of crops - and especially the rate of development (Luo, 2011). Crops grow and develop ideally within the range of optimum temperatures and at a slower rate above or below the range (see Rötter & van de Geijn, 1999).[4] [5] Rötter & van de Geijn (1999), Hatfied (2008) and Luo (2011) provide a review of optimal temperatures for several crops. *Drought stress or excessive moisture* can quickly lead to crop failure or the inability to timely plant or harvest a crop (Rötter & van de Geijn, 1999).

Weather affects indirectly crops due to its effect on associated pests. Precipitation seems to be the most important variable affecting crop-pest interactions (Rosenzweig *et al.*, 2001). Moisture stress on crops make them more vulnerable to be damaged by pests while moisture excess increase the prevalence of fungal pathogens. High temperatures increase pests population.

[4] The relation is not linear, as yield increases between base and optimal temperature but falls between the optimal and failure point temperature (Roberts and Summerfield 1987; Wheeler *et al.* 2000).

[5] Exposure to higher temperatures causes faster development in non-perennial crops, which does not translate into an optimum for maximum production because the shorter life cycle means smaller plants, a shortened reproductive phase duration, and reduced yield potential because of reduced cumulative light interception during the growing season (Hatfield et al. (2011)).

Weather, pesticides and health. To mitigate the effect of excessive rainfall on pests, farmers can increase the use of pesticides. Maternal exposure during pregnancy has been associated with an increased risk of fetal death and spontaneous abortion (Arbuckle and Server, 1998). There is increasing evidence that *in utero* exposure to pesticides increases the risk of growth retardation and low birth weight (Berkowitz *et al.* 2004, Burdorf *et al.* 2011).[6] Breast milk can also be contaminated with pesticides. [7] Susceptibility to infections is suspected but evidence is limited.[8]

Effect on maternal labour supply The effect on maternal labour supply can be ambiguous. The harvesting period is the most labour intensive of the agricultural cycle. If weather reduces agricultural production the demand of labour falls, depressing wages. Given that maternal and child health productions are intensive in time and the value of time falls, mothers can substitute time away from labour market towards health-preserving activities such as prenatal care or breastfeeding. However, the effect will go in an opposite direction if households respond to the negative income shock using maternal labour supply as an insurance mechanism.

This channel seems to be crucial to understand the mixed evidence found in developing countries. Bhalotra (2010) reported recessions increase maternal labour supply in rural areas and infant mortality, where maternal participation in rural agricultural activity has an adverse effect on health seeking behavior. Miller & Urdinola (2010) explored the effect of coffee price shocks in coffee-growing areas in Colombia and found that negative price shocks reduce maternal labour supply and infant mortality. Both reported an inverse relationship between maternal labour supply and health seeking behavior.

Effect on maternal workload. Women in developing countries often continue their agricultural work during late pregnancy. Prematurity has been associated to the increase in agricultural labor during harvesting (Rayco-Solon, Fulford & Prentice, 2005). The epidemiological literature indicates that heavy work is a risk factor for prematurity (Lima, Ashworth & Morris, 1999, Escriba-Aguir *et al.*, 2001 and the review of Bonzini et al. 2007). The fact that much of the work during harvesting involves standing and bending may exacerbate this risk (Hatch et al. 1997), especially bend-

[6]Some pesticides are now suspected of being endocrine disrupting chemicals and have been linked to adverse effects on either embryonic development or reproductive function.

[7]Residues of pesticides have been detected in breast milk (including DDT, HCB and HCH isomers) in contaminated areas (Pronczuk *et al.* 2002).

[8]Inuit infants from the Artic exposed *in utero* and to breast milk contaminated with p,p´-DDE, HCB and dieldrin had an elevated risk of otitis media (Dewailly *et al.* 2000)

ing, for more than one hour a day, during the last month of the pregnancy (Bonzini et al. 2011). This channel suggests that unfavorable weather conditions that reduce agricultural production and maternal workload during harvesting might be beneficial for infant health. Pre-term delivery is a major determinant of perinatal mortality, and of neonatal and infant morbidity.

1.2.2 Weather and health

Extreme weather events such as floods, storms, droughts and heat waves can lead to adverse health outcomes, ranging from physical injuries to heat stress and respiratory diseases. Heat waves cause rash, syncope, cramps, exhaustion, and stroke. Children are less able to control their local climate, especially if a heat wave is sudden and severe (Bunyavanich *et al.*, 2003). Heavy rainfall events can transport terrestrial microbiological agents into drinking-water sources resulting in outbreaks of infectious diseases.

Weather fluctuations can affect the transmission of mosquito-borne diseases (malaria and dengue). Increased rain may increase vector survival by increasing humidity or larval habitat and vector population size by creating new habitat. Excess of rain can eliminate habitat by flooding, decreasing vector population. Low rain can create habitat by causing rivers to dry into pools or increase container-breeding mosquitoes by forcing increased water storage. Temperature affects vector survival, rate of vector population growth, feeding rate and host contact (may alter survival rate) (WHO, 2003).

Pregnant women are the most vulnerable to malaria (Rogerson *et al.* 2007). Malaria causes anemia in pregnant women, premature delivery, fetal growth retardation and low birth weight. [9] Infants exposed to malaria *in utero* have an increased risk of malaria infection and anemia during childhood. Placental infections occurring at the end of pregnancy are those most likely to have an impact on infants' survival (Bardaji *et at.* (2011)).

In this paper I will control for diseases that can affect infant health in the prenatal and postnatal period such as mosquito-borne diseases.

1.3 Empirical Strategy

This section describes the empirical strategy used to analyze the effect of weather-related shocks on infant mortality.

[9] See Desai *et al.* (2007) for a review and Steketee *et al.* (2001).

In order to measure agricultural productivity I define two variables, (i) the fraction of time during the growing season with ideal temperature, and (ii)increased rainfall (rainfall deviation from the parish growing season (October - March)) mean (measured in cm).[10] The first variable measures of the number of months that the monthly mean temperature was in the ideal interval for the main crop in the parish divided by 6 (duration of the growing season). Optimal temperature depends on each stage of the plant's life cycle and is crop specific. Here, I only consider optimum temperature thresholds to the entire growing season. The ideal temperature for each crop is 26-30 Cº (banana), 20-26 Cº (coffee), 21-28 Cº(cocoa), 25-35 Cº(sugar cane), 18-30 Cº (maize), 23-30 Cº (rice), 20-30 Cº (barley) and 15-23 Cº (wheat).[11] Table 17 in the Appendix I summarizes the agriculture cycle.

A basic specification for infant mortality rate, *IMR*, of birth cohort (year-month) j, born in parish p, is:

$$IMR_{jp} = \alpha + \sum_{h=0}^{3} W(t-1)_{jp} * U_{hjp}\beta_h^U + \sum_{k=1}^{4} W(t)_{jp} * L_{k,jp}\beta_k^L + X_{jp}\phi + \lambda_j + \lambda_p + \varepsilon_{jp}, \tag{1}$$

where t and $t-1$ indicates the harvest during the first year after birth and the year before, respectively. $W(t)_{jp}$ is a vector of weather variables that proxy for agricultural productivity at the harvest t for the birth cohort j that was born in parish p, including the fraction of time with ideal temperature and increased rainfall during the growing season. To allow for non linear effects I include quadratic terms for rainfall and temperature. Appropriate rains during the growing period results in good growth of crops but excessive rains can damage them. For each parish p, I compute the value of each crop and I select the main crop based on the highest participation in the total production.[12]

At the time of the harvest a child can be at different stages in the in the fetal period or postnatal period. Table 18 in Appendix I summarizes the stage at the time of the harvest based on the month of birth. According to the mechanisms described in section 1.2, we can expect the magnitude of the effects on infant mortality to vary depending on the stage during the fetal and postnatal period. I interact weather with a dummy variables that indicates the trimester in each stage. Let h be the trimester of pregnancy

[10] The main harvest takes place between April to June (Census of Agriculture 2000).

[11] Temperature intervals are based on Rosenzweig *et al.* (2001), Luo (2011), and Hatfield *et al.* (2008).

[12] The main crop is selected from a list of eight crops - banana, coffee, cocoa, sugar cane, maize, rice, barley and wheat, based on its value. Crop's production at parish level is based on the census of agriculture 2000. Prices are an average from 1995-2009 based on annual national producer prices reported by FAOSTAT.

where $h = 0$ is defined as preconception and $h \in \{0,3\}$. Let k be the trimester after birth where $k \in \{1,4\}$. U_{hjp} are dummy variables that indicate the trimester in the fetal period while L_{kjp} indicate the trimester after birth, for the birth cohort j that was born in parish p. Then β_h^U and β_k^L capture the impact of *within-parish* fluctuations in weather, during the growing season, at the time of the harvest for cohorts that were exposed at different stages of the prenatal and postnatal period, respectively. These coefficients measure the net effect from changes in maternal labour supply and economic resources during the harvest.

X_{jp} is a vector that includes a set of international prices for the main crop in parish p at the time of the harvest during the prenatal and postnatal period where each price is interacted with a set of dummy variables on the stage for the prenatal and postnatal period. It also includes controls for the effect of weather on health during each trimester of the pregnancy and life, for the birth cohort j in parish p.[13] I create a set of variables on excessive rainfall, drought, cold (<12 Cº) and hot (22-28 Cº) temperatures. Each variable reflects the fraction of time during each trimester that an specific climatic event took place. Then weather variables take values in the set $\{0, \frac{1}{3}, \frac{2}{3}, 1\}$. Monthly rainfall is normalized based on the mean and standard deviation for each parish and month. Excessive rainfall is defined as higher than 2 standard deviation. Drought is defined as lower than one standard deviation in order to capture moderate and severe episodes. [14] I allow differential effects of weather on health depending on whether it is a malaria area or not. Malaria ideal temperature for survival and transmission is captured with variables of temperature higher or equal than 22 Cº (hot temperature) and its interaction with malaria areas. [15] The coefficients of the interactions between malaria area and hot temperature will control for the effect of malaria on infant mortality at different stages during the prenatal and postnatal period.

λ_j is a set of birth cohort (or month-year) fixed effect that controls for aggregate shocks; λ_p is a set of parish of birth fixed effect that account for any time-invariant differences between parish that might be correlated with weather and infant mortality, and ε_{jp} is the error term.

Standard errors are clustered at parish level in order to correct for autocorrelation of arbitrary form in shocks to infant mortality across time within a parish.

[13]Transmission of vector-borne diseases are influenced by environmental and climatic factors. Abnormal rainfall and flooding increases availability of breeding places while heavy precipitations and landslides could reduce mosquito abundance if breeding sites are washed away.Temperature accelerates vector development, decreases incubation period and increases feeding frequency.

[14]Normal rain and warm temperature are the base categories.

[15]Barreca (2010) instruments malaria using daily mean temperatures between 22ºC and 28Cº and found that is a strong predictor of malaria death rate.

After establishing the relationship between growing season weather and infant mortality, I will next consider what behavioral responses or mechanisms might explain this results. To study maternal labour supply responses and its effect on time intensive investment decisions I use data from the LCS and RHS surveys, and birth records certificates.

For pregnant women i observed in month m, year y and parish p I estimate

$$y_{imyp} = \alpha + W'_{myp}\beta + Z'_{impy}\theta + \lambda_m + \lambda_y + \lambda_p + \varepsilon_{imyp}, \qquad (2)$$

where y is the number of hours worked during the previous week, W is a vector of the weather during the growing season, Z is a vector that includes individual characteristics (age and education), crops external prices and health-weather related variables, λ_p,λ_m,λ_y are dummies variables for parish, month and year. I modified equation 2 to estimate the probability of work conditional on pregnancy using a probit model. I also explore variation in the month of the interview interacting weather and price variables with the agricultural stage: growing, harvest and post harvest period.

Finally, I used the RHS, LCS, and birth certificates to study time-intensive health investments (prenatal care, place of delivery, breastfeeding and attendance to health check up).

For a children i, born in the cohort j (month-year) in parish p, prenatal care investments are estimated as

$$y_{ijp} = \alpha + \sum_{h=0}^{3} W(g-1)'_{jp} * U_{h,jp}\beta_h^U + Z'_{ijp}\phi + \lambda_j + \lambda_p + \varepsilon_{ijp}, \qquad (3)$$

where y is the number of prenatal care visits or place of delivery (delivery at home instead of health at health facilities), and all other variables are the same than in equation 1 but only for the *in utero* period. Vector Z also includes maternal characteristics as age at birth and education. Breastfeeding and postnatal care are estimated from equation 3 but considering only shocks after birth.

1.4 Data Sources and Summary Statistics

1.4.1 Data Sources

I collected data on infant mortality, temperature, precipitation, agricultural production and prices, malaria and maternal use of health services and child health.

National Vital Records. Birth and death certificates include month and year of each event, parish of residence, maternal characteristics[16], use of health services at birth (skilled attendance), and health measures of the newborn.

Between 1995 and 2008, 1,053,000 children were born in rural Ecuador. I linked those births to deaths records based on maternal place of residence (parish) and date of birth (month and year). For each cohort, infant mortality is the rate of infant death during the first year after live birth, expressed as the number of such deaths per 1000 live births in a parish. [17]

Weather data. I obtained monthly weather data between 1995 and 2009 from the National Institute of Meteorology and Hydrology (INAMHI). Precipitations are available for 340 weather stations while temperatures are restricted to 167 stations across Ecuador (see figure 3). Data include latitude and longitude of each weather station.

To link weather data to infant mortality I used a digital map containing the delimitation of the administrative boundaries down to the 3rd sub national level (parish) provided by the INEC. [18] I compute the distances between the center of each parish and all weather stations. In order to avoid the problem of missing information, I consider the 5 closest weather stations with non missing data for a given month within an radius that varies depending on the weather variable. I choose a radius of 39 km for rainfall and 50 km for temperature.[19] Data from different weather stations may be linked to the same parish over time. I construct parish-month-year weather variables by taking weighted average of the 5 closest weather stations within the radius under analysis. The weights are the inverse of the squared distance from the district center.

Crops Prices. Monthly world prices of crops are reported by the World Bank (Pink Sheet). Annual data on producer prices for Ecuador come from the FAOSTAT database. Prices used in the regressions are expressed in terms to the CPI and detrended.

Agricultural production data. The Census of Agriculture 2000 provides the most comprehensive information on agriculture in Ecuador. It contains data on land use, area planted and harvested, total output, yield per hectare, and agricultural production, among others. This is the only source of data that allows me to work at parish level.

[16]Maternal name has been censored, so it is not possible to match siblings.

[17]Estimated sub-report in birth and deaths is approximately 13% (RHS 2004).

[18]INEC reports historical changes observed in the administrative levels. Based on that information, I updated birth and death locations to the 2009 administrative boundaries.

[19]To choose the radius I selected the 20 closest weather stations for each parish and compute the 95th percentile of the distance distribution.

Agriculture surveys were carried out before 1995 and between 2002 and 2009 but they are only available at province level.

Using the Census 2000, I selected one crop for each parish from a list of eight commodities (banana, coffee, cocoa, sugar cane, maize, rice, barley and wheat) based on the maximum value. I computed the value of the crop as the total output times the average producer price from 1995 to 2008.[20]

Malaria. The spatial distribution of *Plasmodium falciparum* malaria endemicity in Ecuador is reported by Hay *et al.* (2009). I merge the malaria map (see figure 4 in the Appendix I) with the administrative map to identify malaria free parishes.

Household Surveys. Reproductive Health Survey (RHS) 1999 and 2004 are representative for urban and rural areas. They collect information about use of health care services, and current and retrospective behavior about prenatal care and breastfeeding. I pooled the 1999 and 2004 to have information for children born between 1996 and 2004.

Three surveys (1998, 1999 and 2006) of the Living Condition Survey were carried out during the period of interest. They collect information about labor, income and use of health services and health behavior. They also report expenditures on pesticides.

Data Limitations. The agronomic literature details the importance of having daily data to measure crops' ideal temperature at different stages. Then monthly data, as I consider, might underestimate any effect.

It is important to note that crops production is not available at parish level (except for the census 2000). Then it is not possible to include agricultural productivity as an endogenous variable in equation 1 and instrument it with growing season weather.

1.4.2 Summary Statistics

Simple average in rural areas shows that infant mortality rates are higher for children born during the first semester of the year (figure 1). Figure 2 reports the average of total rural birth between 1996 to 2008. Births are well distributed during the year with a small reduction in the last trimester of the year that corresponds to the beginning of the rainy season and the first trimester of the growing season. [21]

[20]For coffee, I used the producer price in Ecuador reported by the International Coffee Organization.
[21]In section 1.6 I will show my results are not connected with fertility decisions.

FIGURE 1
Infant Mortality Rate

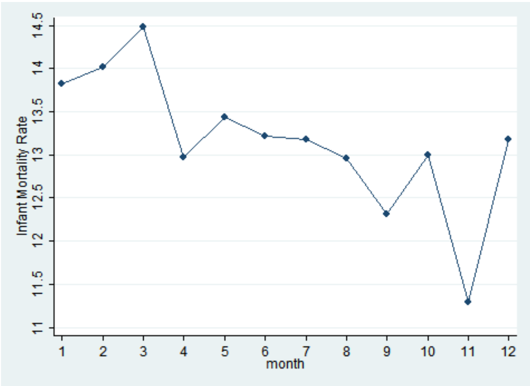


FIGURE 2
Rural births 1996-2008 by month

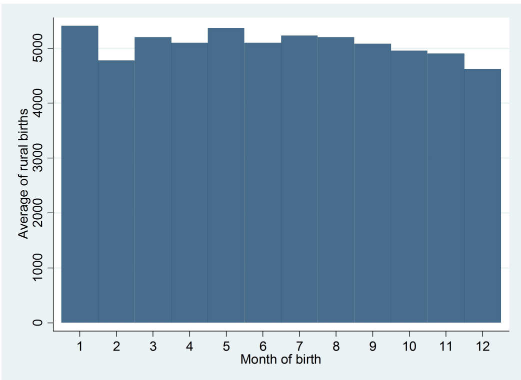


TABLE 1

Marginal effect of agricultural productivity on infant mortality rate

Children born between 1996 to 2008	Time of the Harvest							
	In Utero				After Birth			
	Pre C	1st Trim	2nd Trim	3rd Trim	1st Trim	2nd Trim	3rd Trim	4th Trim
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ideal Temperature - fraction of time	2.647	19.571***	3.169	7.840	4.892	-11.376*	1.328	3.168
	(6.645)	(6.808)	(6.073)	(6.738)	(6.586)	(6.213)	(6.398)	(6.621)
Increased rainfall x 10	-0.105	0.130	-0.059	0.694***	0.481*	-0.004	0.186	0.329
	(0.195)	(0.248)	(0.212)	(0.237)	(0.290)	(0.254)	(0.297)	(0.220)
N = 63,398 Mean=15.9								
Implied Effect on IMR								
Ideal Temperature - fraction of time (1 pp)	-	1.2%	-	-	-	-0.7%	-	-
Increased rainfall	-	-	-	0.4%	0.3%	-	-	-

Note: Ideal temperature is a measure of the fraction of time during the growing season that the temperature for the main crop in the parish falls in the optimal interval. Increased rainfall is defined as the parish growing season rain minus its mean. Marginal effect on rainfall is reported at one standard deviation. Regression controls for interactions between international price and time of the harvest during the first year after birth and the previous year, weather health-related variables during pregnancy and first year after birth, cohort fixed effect (time fixed effect - monthly), and parish fixed effect. The implied effect is computed considering the mean of infant mortality rate (15.9 per 1000 live births). Standard errors are clustered at the parish level and reported in parenthesis. * Significant at 10%, ** significant at 5%, *** significant at 1%.

1.5 Results

This section reports the results of agricultural productivity proxy variables on infant mortality rate based on equation (1). Table 1 displays the marginal effects of deviation of rainfall from the parish average, and the fraction of time the temperature falls in the optimal interval for the main crop in the parish, both during the growing season.

Unfavorable weather conditions during the growing season reduces infant mortality when the harvest takes place during the the first and third trimester of the *in utero* period and during the first trimester of life. 1 percentage point (pp) decrease in the fraction of time with ideal temperature is associated with a 0.19 reduction in infant mortality per 1000 children born. [22] Since the average IMR in the sample is 15.9 per 1000, the implied effect on IMR is a reduction by 1.2%. The marginal effect of growing season rainfall on infant morality rate, when rainfall is one standard deviation below the mean, is -0.07 per 1000 (implying a reduction by 0.4%) for a cohort exposed to the harvest in the third trimester of the *in utero* period. The effect is -0.05 per 1000 (-0.3%) when the harvest occurs during the neonatal period (first trimester after birth).

[22] Ideal temperature is a count of the number of months during the growing season that the temperature was in the optimal interval. Given the temperature in each month has a different distribution, then one standard deviation in the ideal temperature will not informative about observed deviations in temperature in each month. To consider a relevant measure for the change in the marginal effect, I compute the coefficient of variation for temperature (0.085) based on the mean (19.11 Cº) and the standard deviation (1.64 Cº) by month and parish. Hence, temperatures deviate from the mean by 8.5%. The variable of the interaction between ideal temperature during the growing season and first trimester of pregnancy at the time of the harvest has a mean value of 0.10, then assuming that ideal temperature deviate in 8.5% that will be an increment in 0.85pp. To simplify the interpretation I will consider 10% deviation in ideal temperature from the mean (0.10) and use 1pp increment to report the results.

TABLE 2

Marginal Effects of international prices and weather-health variables on IMR

Children born between 1996 - 2008

	In Utero				After Birth			
	Pre C	1st Trim	2nd Trim	3rd Trim	1st Trim	2nd Trim	3rd Trim	4th Trim
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
International Prices	-0.129	-0.405**	0.110	0.001	0.048	-0.117	0.085	-0.160
	(0.175)	(0.162)	(0.111)	(0.069)	(0.073)	(0.074)	(0.267)	(0.233)
MALARIA FREE AREA								
Rain excess (> 2 s.d.)		-3.398	-0.226	3.406	1.139	4.976	-2.211	3.159
		(2.854)	(3.102)	(3.195)	(3.047)	(3.504)	(2.763)	(2.631)
Drought (< -1 s.d.)		-0.295	-0.419	-0.293	1.115	1.884	4.137**	1.227
		(1.840)	(1.937)	(1.808)	(2.088)	(1.887)	(2.018)	(1.903)
Temperature <12 C°		-0.458	5.437*	-1.221	-5.438*	3.315	2.592	-3.487
		(2.688)	(2.876)	(2.798)	(2.797)	(3.425)	(3.332)	(2.705)
Temperature 22-29 C°		3.083	-4.571*	2.354	3.217	-2.410	-0.654	-0.219
		(2.163)	(2.552)	(2.424)	(2.795)	(2.756)	(2.827)	(2.630)
MALARIA AREA								
Rain excess (> 2 s.d.)		-3.520	-0.999	-6.842*	3.190	0.748	4.077	-4.187
		(3.058)	(3.952)	(4.091)	(3.641)	(3.766)	(3.472)	(3.048)
Drought (< -1 s.d.)		1.149	5.848*	-1.024	2.161	-0.688	1.723	-3.976
		(3.191)	(3.234)	(2.737)	(3.334)	(3.130)	(3.419)	(2.909)
Temperature <12 C°		17.924	-19.951	-4.597	-0.885	-5.236	7.193	0.009
		(19.823)	(16.312)	(6.001)	(5.801)	(6.820)	(5.358)	(6.620)
Temperature 22-29 C°		-0.277	-0.708	-2.648	2.688	-2.430	2.601	-1.314
		(2.984)	(2.691)	(2.928)	(3.287)	(4.099)	(4.245)	(3.032)
Interactions								
Rain excess x Malaria Area		-0.122	-0.773	-10.247**	2.051	-4.228	6.288	-7.348*
		(3.801)	(4.877)	(5.071)	(4.448)	(4.875)	(4.415)	(3.950)
Drought x Malaria Area		1.445	6.267*	-0.731	1.047	-2.572	-2.414	-5.203
		(3.617)	(3.571)	(3.166)	(3.619)	(3.422)	(3.714)	(3.334)
Temperature <12 C° x Malaria Area		18.381	-25.389	-3.376	4.553	-8.550	4.601	3.495
		(20.035)	(16.544)	(6.622)	(6.315)	(7.577)	(6.293)	(7.096)
Temperature 22-29 C° x Malaria Area		-3.360	3.862	-5.002	-0.529	-0.020	3.254	-1.095
		(3.583)	(3.614)	(3.765)	(4.263)	(4.942)	(5.043)	(4.007)

Note: Sample size 63,398. Standard errors are clustered at the parish level and reported in parenthesis. * Significant at 10%, ** significant at 5%, *** significant at 1%.

I define the neonatal period as the first trimester after birth[23]. The beneficial effect of unfavorable weather conditions reverts after the first trimester of life. If the harvest falls during the second trimester after birth, a reduction by 1pp in the fraction of time with ideal temperature increases the IMR by 0.7% (0.11 per thousand).

Exposure to agricultural shocks generates relatively larger substitution effects during the pregnancy and the neonatal period. The net effect seems to be stronger at the beginning of the pregnancy. Fetal health is more vulnerable during the first trimester of the pregnancy and shocks that occur during that time are more likely to lead to fetal death. Then, we might expect to underestimate the effect of shocks during the first trimester of the pregnancy.

[23] Note that the neonatal period is usually defined as the first 28 days after birth. Here I'm extending it to the first trimester of life.

TABLE 3

Marginal effects of agricultural productivity on IMR - Malaria free area

Children born between 1996 to 2008

	In Utero				After Birth			
	Pre C	1st Trim	2nd Trim	3rd Trim	1st Trim	2nd Trim	3rd Trim	4th Trim
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ideal Temperature - fraction of time	-0.045	26.797***	3.725	10.991	8.020	-7.158	1.991	0.049
	(8.867)	(9.476)	(8.011)	(9.087)	(8.384)	(8.319)	(8.308)	(8.383)
Increased rainfall x 10	-0.025	0.155	0.842*	0.579	0.091	-0.073	-0.444	-0.464
	(0.414)	(0.556)	(0.434)	(0.362)	(0.417)	(0.439)	(0.429)	(0.382)
N= 43,100								
Implied Effect on IMR								
Ideal Temperature - fraction of time (1 pp)	-	1.7%	-	-	-	-0.5%	-	-
Increased rainfall	-	-	0.5%	0.4%	-	-	-	-

Note: Ideal temperature is a measure of the fraction of time during the growing season that the temperature for the main crop in the parish falls in the optimal interval. Increased rainfall is defined as the parish growing season rain minus its mean. Marginal effects on rainfall are reported at one standard deviation. Regression controls for interactions between international price and time of the harvest during the first year after birth and the previous year, weather health-related variables during pregnancy and first year after birth, cohort fixed effect (time fixed effect - monthly), and parish fixed effect. The implied effect is computed considering the mean of infant mortality rate (15.9 per 1000 live births). Standard errors are clustered at the parish level and reported in parenthesis. * Significant at 10%, ** significant at 5%, *** significant at 1%.

Table 2 reports the effect of control variables (international prices and health related weather) on infant mortality rate related to table 1. Increments in detrended relative prices have a beneficial effect on mortality when they occur at the beginning of the pregnancy (income effect is larger than the substitution effect). The effect on infant mortality of 10pp increase in prices is -0.0405 per 1000 (-0.25%). Increments in temperature during the second trimester of pregnancy are related with a reduction in mortality (table 2). Temperature is a measure of the number of month with hot and cold mean temperature. It captures the general effect of hot and cold temperature relative to warm weather but might fail to reflect isolated events such as one very hot day in the month. Excess of rain at the end of the pregnancy has a beneficial effect on infant mortality in malaria areas while a drought has the opposite effect. Heavy rainfalls and high water levels can wash away mosquito habitats. In very humid climates, droughts may turn rivers into strings of pools that provide good breeding sites for mosquitoes.

To explore whether the effects reported in table 1 are driven by malaria areas I estimate equation (1) separately for parish where malaria is endemic and those where it is not. Table 3 and 4 report the marginal effects of agricultural productivity on IMR in areas free of malaria and areas where malaria is endemic, respectively. The effects of ideal temperature during the growing season for shocks that take place during the pregnancy are positive in both areas but higher in malaria free areas. For the third trimester of pregnancy the effect of rains is similar in both areas. When harvest occurs during the neonatal period the effect of growing season rainfall is stronger in malaria areas. These findings confirm that the results during the *in utero* period are not driven by the effect of weather on vector born diseases.

TABLE 4
Marginal effects of agricultural productivity on IMR - Malaria area

Children born between 1996 to 2008	In Utero				After Birth			
	Pre C	1st Trim	2nd Trim	3rd Trim	1st Trim	2nd Trim	3rd Trim	4th Trim
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ideal Temperature - fraction of time	3.942	7.285	1.968	2.086	-5.746	-16.340*	-3.339	6.296
	(9.051)	(9.266)	(8.852)	(9.075)	(10.756)	(9.078)	(10.231)	(10.565)
Increased rainfall x 10	0.011	-0.149	-0.558*	0.517	0.731*	-0.060	0.427	0.706
	(0.240)	(0.330)	(0.329)	(0.318)	(0.375)	(0.392)	(0.399)	(0.297)
N= 20,298								
Implied Effect on IMR								
Ideal Temperature - fraction of time (1 pp)	-	0.5%	-	-	-	-1.0%	-	-
Increased rainfall	-	-	-0.4%	0.3%	0.5%	-	-	-

Note: Ideal temperature is a measure of the fraction of time during the growing season that the temperature for the main crop in the parish falls in the optimal interval. Increased rainfall is defined as the parish growing season rain minus its mean. Marginal effects on rainfall are reported at one standard deviation. Regression controls for interactions between international price and time of the harvest during the first year after birth and the previous year, weather health-related variables during pregnancy and first year after birth, cohort fixed effect (time fixed effect - monthly), and parish fixed effect. The implied effect is computed considering the mean of infant mortality rate (15.9 per 1000 live births). Standard errors are clustered at the parish level and reported in parenthesis. * Significant at 10%, ** significant at 5%, *** significant at 1%.

1.6 Mechanisms

In this section I investigate the effect of weather on maternal labour supply and the impact through this channel. I hypothesize that unfavorable weather conditions reduces crops yield, the probability of work, and hours worked during the harvest, affecting time investment in child health. I also show that agricultural productivity does not affect fertility.

Maternal labour supply

About 70% of rural women aged 15-49 work in Ecuador. I investigate the labor supply and labor income of pregnant women in rural areas using equation 2. Information on work participation is available at the time of the interview, dated on November 1998 until October 1999, and November 2005 until November 2006.[24] Growing season is defined from October until March. I define year t from October $t-1$ until September t. November 2006 was excluded to report full years.

Table 5 indicates that increments in crop's ideal temperature and increased rainfall during the growing season increases hours worked, probability of work and labor income for pregnant women. 1 pp increase in the fraction of time with ideal temperature increases hours worked of pregnant women by 0.50 hours (3%) while one standard deviation of rainfall increases it by 1.2 hours (3%).

To explore the effect during the harvest, I interact weather variables and prices with dummies for growing, harvest and post-harvest period. In this case, quadratic terms

[24] LSMS 1998 is excluded because the interview date is not reported in the publicly available database.

TABLE 5

Labor supply and labor income of pregnant women

	OLS Hours Worked	Probit Prob Work	OLS log Income
	(1)	(2)	(3)
Ideal Temperature - fraction of time	50.202***	4.941*	11.522**
	(17.451)	(2.570)	(4.966)
Increased rainfall x 10	11.816**	5.405	3.935*
	(5.754)	(6.641)	(2.098)
Mean	18.8	0.51	18.8
N	244	85	244
Implied Effect			
Ideal Temperature - fraction of time (1pp)	3%	10%	7%
Increased rainfall	6%	-	23%

Note: Based on LCS 1999 and 2006. All regressions include controls for age, education, dummies for month, year, parish, and health controls for hot and cold temperature, excess of rain and drought for the last month and interactions with malaria areas. Dependent variables are hours worked in the last week, work last week, log (income +1) expressed in dollars at december 2006. Standard errors are clustered at the parish level and reported in parenthesis. * Significant at 10%, ** significant at 5%, *** significant at 1%.

for weather and prices variables were excluded. Table 6 suggests that an increment in ideal temperature increases the hours worked during the harvest, while increased rainfall increase hours worked in all periods. Higher prices during the harvest reduce the number of hours worked for pregnant women.

Prenatal Care

The use of prenatal care services is time intensive, involving travel and waiting times. It improves birth outcomes and newborn survival. During the prenatal care visits pregnant women receive tetanus immunization, malaria prevention and treatment, man-

TABLE 6

Hours worked of pregnant women

	Hours Worked		
	Growing	Harvest	Post-Harvest
	(1)	(2)	(3)
Ideal Temperature - fraction of time	29.437	68.043*	-23.409
	(36.612)	(39.461)	(21.502)
Increased rainfall x 10	14.440***	13.354**	20.548***
	(4.968)	(7.276)	(7.778)
N= 244			
Implied Effect			
Ideal Temperature - fraction of time (1pp)		4%	
Increased rainfall	8%	7%	11%

agement of anemia, and treatment of STIs that can improve fetal and child health outcomes. Mothers at risk of delivering a preterm or growth-retarded infants are also identified. Prenatal care is the link to promote the benefits of skilled attendance at birth and to encourage women to seek postnatal care for their newborns and themselves.

TABLE 7

Marginal Effect on Prenatal Care

Children born between 1996 to 2006

	Time of the harvest: In Utero			
	Pre C	1st Trim	2nd Trim	3rd Trim
	(1)	(2)	(3)	(4)
Ideal Temperature - fraction of time	-1.3648	0.0386	0.5860	0.3489
	1.224	1.266	1.360	1.380
Increased rainfall x 10	-0.0001	-0.0232	0.0217	-0.1221**
	0.036	0.038	0.038	0.048
N=3,295 Mean=3.8				

Note: The dependent variable is the number of prenatal care visits. The mean value is 3.8 visits. Marginal effect on rainfall is reported at one standard deviation. Regression controls for mother education and age at delivery, interactions between international price and time of the harvest during the year before birth, weather health-related variables during pregnancy, birth cohort fixed effect, and parish fixed effect. Standard errors are clustered at the parish level and reported in parenthesis. * Significant at 10%, ** significant at 5%, *** significant at 1%.

In this section I explore the effect agricultural productivity shocks on the frequency of use of prenatal care and no skilled attendance at delivery (delivery at home) using equation 3 and OLS and linear probability models, respectively.

During the third trimester of pregnancy, prenatal care visits should be more frequent. Biweekly from 28 to week 36 of pregnancy and weekly after week 36. To evaluate the effect on prenatal care visits I consider the stage of the pregnancy at the time of the harvest and whether it was a good or bad harvest season based on weather conditions during the growing season. My results indicate that the use of health services falls when the harvest takes place during the third trimester of the pregnancy and there was a favorable weather during the growing season (tables 7 and 8). The number of prenatal care visits decreases by 1.22pp (0.3%) and the probability of delivering the child at home increases by a range of 0.02-0.17 pp (0.03% - 0.4%).

Breastfeeding and Postnatal Care

Breastfeeding is a time-intensive investment that plays a crucial role in child survival. Maternal milk provides all the nutrients required for a healthy growth and development of newborns, specially during the first 6 months, and protects children from diarrhea

TABLE 8

Marginal Effect on Place of Delivery

Children born between 1996 to 2006

	Time of the harvest: In Utero			
	Pre C	1st Trim	2nd Trim	3rd Trim
	(1)	(2)	(3)	(4)
A. LIVING CONDITIONS SURVEY (1998, 1999 & 2006)				
Ideal Temperature - fraction of time	0.0112	-0.2229	-0.0827	-0.1704
	(0.2536)	(0.2479)	(0.2681)	(0.2440)
Increased rainfall x 10	0.0069	-0.0148	0.0081	0.0166**
	(0.0078)	(0.0088)	(0.0087)	(0.0074)
N=1,910 Mean=0.42				
B. BIRTH CERTIFICATES (1996-2008)				
Ideal Temperature - fraction of time	0.0054	0.0010	0.0186	0.0197
	(0.0180)	(0.0210)	(0.0219)	(0.0205)
Increased rainfall x 10	-0.0005	0.0003	0.0009	0.0019***
	(0.0007)	(0.0006)	(0.0007)	(0.0007)
N=456,366 Mean=0.56				

Note: The dependent variable is an indicator for delivery at home. Marginal effects are reported from LPM estimates. Marginal effects on rainfall are reported at one standard deviation. All regression controls for mother education and age at delivery, interactions between international price and time of the harvest during the year before birth, weather health-related variables during pregnancy, birth cohort fixed effect, and parish fixed effect. Standard errors are clustered at the parish level and reported in parenthesis. * Significant at 10%, ** significant at 5%, *** significant at 1%.

and respiratory infections that are one of the main causes of infant mortality. I explore times breastfeed in the last 24 hours for children younger than 12 months, and duration of breastfeeding for children between 24 to 48 months that were weaned. Table 9 reports the marginal effects of agricultural productivity on breastfeeding from OLS estimates. Children born at the time of the harvest, when the growing season weather was adverse, are benefited from extra breastfeeding. Rainfalls one standard deviation below the mean increase frequency during the last 24 hours by 4 times for children under 12 months. 1pp fall in crops ideal temperature increases the total duration of breastfeeding by 0.1 months (0.8%). The use of preventive health services for children younger than 12 months doesn't seem to be affected by shocks during the first year of life (table 10).

Fertility

To check the effect of agricultural productivity on the size of the cohort I estimate equation 1 for the year before birth. The dependent variable is the logarithm of the number of births. Table 11 reports no significant effect on fertility.

TABLE 9

Marginal Effects of agricultural productivity on breastfeeding

Children born between 1996 - 2004

	Time of the harvest: After Birth			
	1st Trim	2nd Trim	3rd Trim	4th Trim
	(1)	(2)	(3)	(4)
A. TIMES BREASTFEED (currently breastfeeding) - Children <12 months. RHS (2004)				
Ideal Temperature - fraction of time	-3.472	-30.187	-50.960	-54.501
	(73.128)	(77.832)	(71.434)	(53.998)
Increased rainfall x 10	-40.961*	-23.612	-15.486	-4.784
	(22.777)	(26.033)	(22.159)	(13.646)
N=227 Mean=10.4 times				
B. BREASTFEEDING DURATION (weaned) - Children 24-48 months. RHS (1999 & 2004)				
Ideal Temperature - fraction of time	-11.574**	-3.536	2.774	-5.824
	(4.916)	(5.293)	(7.592)	(6.110)
Increased rainfall x 10	0.991	0.691	0.024	0.198
	(1.174)	(1.010)	(0.522)	(0.278)
N = 1,345 Mean=10.8 months				

Note: Marginal effects on rainfall are reported at one standard deviation. All regressions include child's age and gender, mother's age and education, interactions between international price and time of the harvest during the first year after birth and the previous year, weather health-related variables during pregnancy and first year after birth, cohort (month-year) fixed effect, parish fixed effect. Times breastfeed is defined as the number of times the child was breastfeed between the 6.00 am of the previous day and 6.00 am of the interview day (mean 10.4 times). Breastfeeding duration is defined as the number of month the child has been breastfeed for those children that were weaned (mean 10.8 months). Standard errors are clustered at the parish level and reported in parenthesis. * Significant at 10%, ** significant at 5%, *** significant at 1%.

TABLE 10

Preventive care

Children born between 1996 - 2006

	Time of the harvest: After Birth			
	1st Trim	2nd Trim	3rd Trim	4th Trim
	(1)	(2)	(3)	(4)
Children <12 months. LCS (1998, 1999 & 2006)				
Ideal Temperature - fraction of time	-0.119	0.155	0.018	0.489
	(0.787)	(0.884)	(0.870)	(1.041)
Increased Rainfall x 10	0.029	0.032	0.030	-0.024
	(0.021)	(0.032)	(0.028)	(0.023)
N=569 Mean=0.26				

Note: Dependent variable is a dummy variable on preventive care utilization. Marginal effects are based on linear probability models. Effects of rainfall are reported at one standard deviation. All regressions include child's age and gender, mother's age and education, interactions between international price and time of the harvest during the first year after birth and the previous year, weather health-related variables during pregnancy and first year after birth, cohort (month-year) fixed effect, parish fixed effect.

TABLE 11

Marginal Effect of Agricultural Productivity on Fertility

Children born between 1996 to 2008

	Time of the harvest: In Utero			
	Pre C	1st Trim	2nd Trim	3rd Trim
	(1)	(2)	(3)	(4)
Ideal Temperature - fraction of time	0.039	0.053	0.025	0.064
	(0.054)	(0.051)	(0.048)	(0.054)
Increased rainfall x10	0.002	0.002	-0.001	-0.001
	(0.002)	(0.002)	(0.002)	(0.002)

Note: The dependent variable is the logarithm of the number of births. Sample size is 63,398. Regression includes controls for the interactions between international price and time of the harvest, weather health-related variables during the pregnancy, cohort fixed effect (time fixed effect - monthly), and parish fixed effect. Standard errors are clustered at the parish level and reported in parenthesis. * Significant at 10%, ** significant at 5%, *** significant at 1%.

1.7 Robustness Check

Proxy variables

The Agricultural Surveys (2002-2009) report production at province level and there is no other information at parish level but the census 2000. Although I'm unable to test whether favorable weather conditions for the main crop increases crop yield at parish level, I present the results aggregated at province level. The main crop at province level is defined as the crop with the highest frequency of being the main crop at parish level. Ideal temperature is based on the main crop at province level. Table 12 suggests that increments in ideal temperature and rainfall increase crop yield. This relation is only significant for increased rainfall.

TABLE 12

Crops yield

	Production per harvested area (Tn/ha)	Production per area (Tn/ha)
	(1)	(2)
Ideal Temperature - fraction of time	0.914	0.923
	(0.946)	(1.020)
Increased rainfall x 10	0.604*	0.621*
	(0.344)	(0.348)
N	137	113

Note: Dependent variables represent annual crops yield at province level. Based on the agricultural census (2000) and agricultural surveys (ESPAC 2002-2009). Weather variables during the growing season have been aggregated at the province level. All regressions include lineal and quadratic term of weather variables, year and province dummies variables. Standard errors are clustered at the province level and reported in parenthesis. * Significant at 10%, ** significant at 5%, *** significant at 1%.

Local time trends

Including parish-specific time trends (644 variables) to equation (1) doesn't modify the main conclusion. It only reduces the implied effect for shocks during the first trimester of pregnancy from 1.2% to 0.9% and increases the absolute effect during the second trimester of life from -0.7 to -1% (see tables 1 and 13)

TABLE 13

Marginal effect of agricultural productivity on IMR including local trends

Children born between 1996 to 2008

	Time of the Harvest							
	In Utero				After Birth			
	Pre C	1st Trim	2nd Trim	3rd Trim	1st Trim	2nd Trim	3rd Trim	4th Trim
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ideal Temperature - fraction of time	-1.929	**13.991***	-2.147	2.468	-0.385	**-16.297*****	-3.811	-0.795
	(6.967)	(7.141)	(6.533)	(6.856)	(6.792)	(6.351)	(6.312)	(6.895)
Increased rainfall x 10	-0.049	0.171	-0.024	**0.711*****	0.490	0.006	0.187	0.344
	(0.205)	(0.263)	(0.224)	(0.248)	(0.301)	(0.261)	(0.311)	(0.232)
N = 63,398								
Implied Effect on IMR								
Ideal Temperature - fraction of time (1 pp)	-	**0.9%**	-	-	-	**-1.0%**	-	-
Increased rainfall	-	-	-	**0.4%**	0.3%	-	-	-

Note: Increased rainfall is defined as the parish growing season rain minus its mean. Marginal effect on rainfall is reported at one standard deviation. Regression controls for interactions between international price and time of the harvest during the first year after birth and the previous year, weather health-related variables during pregnancy and first year after birth, cohort fixed effect (time fixed effect - monthly), parish fixed effect, and parish trends. The implied effect is computed considering the mean of infant mortality rate (15.9 per 1000 live births). Standard errors are clustered at the parish level and reported in parenthesis. * Significant at 10%, ** significant at 5%, *** significant at 1%.

El niño event

Several climatic phenomena took place during 1995-2009, episodes of *El Niño Southern Oscillation* (ENSO) occur every 2 to 7 years. During our period took place on 1997, 2002, 2004 and 2006 and *events of La Niña* on 1995, 1999 and 2007. In Ecuador, ENSO episodes are associated with heavy rainfall and increments in the temperature while *la niña* episodes are associated with drought. The episode during 1997 was one of the strongest episodes in the 20th century and in some areas causes flooding and destroys crops. In this section I replicate table 1 excluding the growing season 97/98 and all births between October 1997 to September 1999. I find that the effects during the prenatal period do not change but the effect on the trimester of birth is not observed (see table 14).

TABLE 14

Marginal effect of agricultural productivity on infant mortality rate excluding ENSO episode.

Children born between 1996 to 2008	Time of the Harvest							
	In Utero				After Birth			
	Pre C	1st Trim	2nd Trim	3rd Trim	1st Trim	2nd Trim	3rd Trim	4th Trim
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ideal Temperature - fraction of time	1.259	26.510***	1.048	12.292*	7.191	-15.548**	1.740	-1.799
	(7.358)	(8.405)	(6.701)	(7.428)	(7.421)	(7.756)	(7.050)	(7.413)
Increased rainfall x10	-6.591	5.365	-0.914	6.837***	-3.191	-1.465	1.794	2.395
	(0.834)	(1.255)	(0.205)	(0.241)	(0.638)	(0.800)	(0.659)	(0.284)
N=53371 Mean=15.6								
Implied Effect on IMR								
Ideal Temperature - fraction of time (1 pp)	-	1.7%	-	-	-	-1.0%	-	-
Increased rainfall	-	-	-	0.4%		-	-	-

Note: Excludes the period between October 1997 until September 1999. Standard errors are clustered at the parish level and reported in parenthesis. * Significant at 10%, ** significant at 5%, *** significant at 1%.

Pesticides

To mitigate the effect of weather farmers can increase the use of pesticides as described in section 1.2.1. As discussed, pesticides might affect *in utero* and after birth infant health. Table 15 reports the effect of weather on annual household spending on pesticides. There is no significant effect of increased rainfall on pesticide household expenditure. When interacting agricultural productivity variables with the agricultural time (growing, harvest, post-harvest) I find that favorable temperature during the growing season seems to reduce spending of pesticides during the harvest. This could cause a protective effect on infant health but doesn't seem to be the dominant mechanism.

TABLE 15

Marginal effect of agricultural productivity on pesticide expenditure

	log pesticides (1)
Ideal Temperature - fraction of time	-0.528
	(0.604)
Increased rainfall x 10	-0.333
	(0.283)
N= 1411	

Note: Regressions includes controls for main international price, education and age of household's head, dummies for month, year, parish, and malaria area, monthly health controls for hot and cold temperature, excess of rain and drought during the last 2 months and interactions with malaria areas. Standard errors clustered at parish level (222 clusters) and reported in parenthesis. *** p<0.01, ** p<0.05, * p<0.1. LSMS 1999 and 2006.

TABLE 16

Marginal effect of agricultural productivity on pesticide expenditure by agricultural time

	log pesticides		
	Growing	Harvest	Post-Harvest
	(1)	(2)	(3)
Ideal Temperature - fraction of time	-0.196	-1.775*	-1.307
	(0.747)	(1.064)	(0.962)
Increased rainfall x 10	-0.110	-0.280	-0.190
	-(0.190)	-(0.220)	-(0.130)
N= 1411			

Note: Regressions includes controls for education and age of household's head, dummies for month, year, parish, and malaria area, monthly health controls for hot and cold temperature, excess of rain and drought during the last 2 months and interactions with malaria areas. Regression considers only linear variables for growing season weather. Standard errors clustered at parish level (222 clusters) and reported in parenthesis. *** p<0.01, ** p<0.05, *

1.8 Conclusions

In rural Ecuador 55% of infant deaths occur during the first trimester of life and most of these deaths take place during the the first month of life, a period which is associated to the events that occurs during the pregnancy.

This paper uses weather fluctuations to explore the effect of unanticipated temporary income fluctuations on infant mortality for children younger than 12 months. I find a pro-cyclical relationship between infant mortality and weather conditions during the growing season. Unfavorable growing season weather reduces infant mortality when the harvest takes place during the first and third trimester of the *in utero* period, and the first trimester of life. In particular, 1 percentage point increment in the fraction of time with crops' ideal temperature and one standard deviation in rainfall during the growing season lead to a reduction in infant mortality ranging from 0.4% to 1.2% of its mean.

My results suggest that negative agricultural shocks during the prenatal and neonatal periods have a positive effect on child health. I find evidence that the substitution effect of shocks during the prenatal and neonatal period outweigh the income effect. The main behavioral mechanism underpinning these effects seems to operate through maternal labour supply. The effect of these shocks seems to work more through prenatal and neonatal time-investment in health and nutrition (prenatal care, delivery practices, breastfeeding) than postnatal investments in nutrition and child healthcare.

Appendix I

Table 17 shows the main agricultural cycle. Planting and growing season starts in October and last until the beginning of April when the harvest starts. The main harvest season is between April and June and can be extended until September. I restrict the harvest period from April to June.

TABLE 17
Agriculture Calendar

GROWING	HARVEST	POST-HARVEST
Oct-Mar	Apr-Jun	Jul-Sep

Note: Based on Census of Agriculture (2000)

TABLE 18
Birth cohort and age at the harvest

Birth month/year	Age (months) at the harvest (t)	In Utero stage harvest (t-1)
Apr-Jun (t)	0-2	Pre-conception
Jan-Mar (t)	3-5	Pregnancy (1st trim)
Oct-Dec (t-1)	6-8	Pregnancy (2nd trim)
Jul-Sep (t-1)	9-11	Pregnancy (3rd trim)

FIGURE 3

Weather Stations

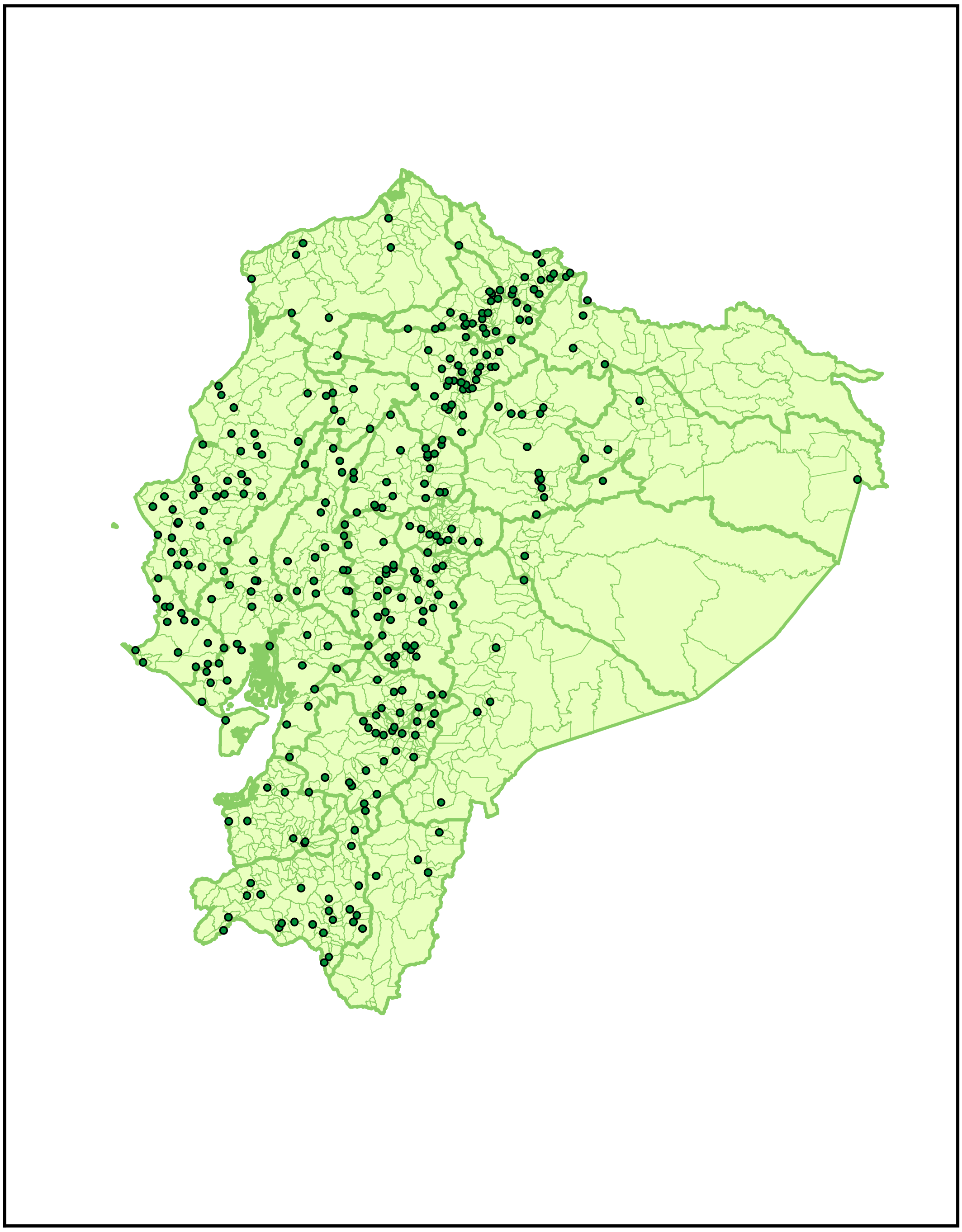
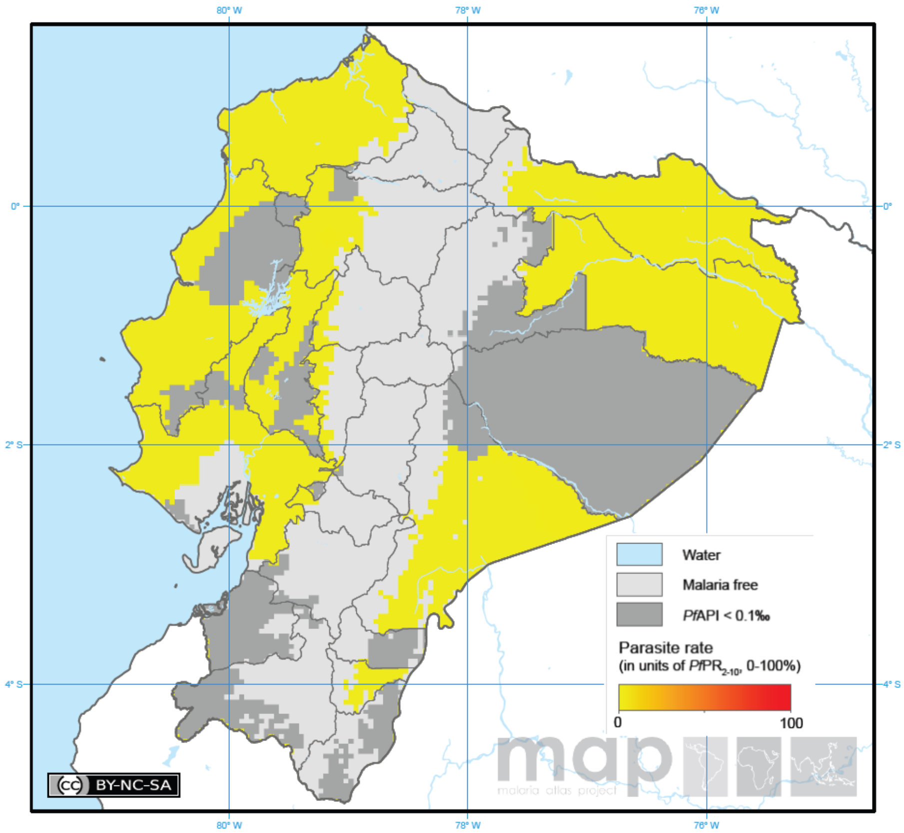


FIGURE 4
Malaria Map



2 Spillover Effects of Conditional Cash Transfers on Schooling Outcomes

Abstract

This paper explores the presence of spillover effects on schooling outcomes from the Colombian welfare program, "Familias en Acción", on ineligible households in rural areas. The program provides cash subsidies to poor families conditional on children school attendance. The identification strategy exploits the fact that exposure to the program varies by locality of residence and date of birth. Using a 10% census sample from 2005 I find positive spillover effects in areas where the program was available. Ineligible children — those living in a household that has not been classified as poor—residing in targeted areas are more likely to stay in the school during the transition period between primary and secondary school. My estimates report an increment by 9 percentage points in primary school completion rate among non-poor children and 7-10 percentage points for grade completion in the initial years of secondary school. Using the partial population experiment setting created by the intervention I extend the analysis to assess the response of the ineligible children to the introduction of the program in their peer group—children from the same birth cohort living in the same municipality. The results suggest that behavioral social interactions (*peer effect*) might play an important role in schooling decisions where the increased grade completion rate of the peer group increases the individual completion. This finding imply that policies aimed at encouraging enrollment can produce large social multiplier effects.

2.1 Introduction

Policy interventions in developing countries are likely to affect other residents in the areas where they are implemented. The program evaluation literature has focused on estimating the impact of the program on the treated or eligible population. In most of the cases, data collection for program evaluation purposes is limited to eligible households restricting the possibility to explore spillover effects of the program on ineligible households. This paper will use data from the Colombian census (2005) to investigate the spillover effects from a large scale conditional cash transfer program introduced in 2001.

The program "Familias en Acción" is a large scale welfare program that provides grants to poor families in Colombia contingent to school attendance and children visits to health services. The program is targeted on poor families based on a poverty score called SISBEN and an eligibility threshold. Findings from the evaluation of the program in rural areas reported an increment in the enrollment rate by 7 percentage points on poor children in secondary school age (Attanasio *et al.* 2010) and increased consumption of poor families by 19.5% (Attanasio *et al.* 2005). The intervention and the positive effects found on poor households might affect ineligible (noon poor) households in an a priori non-trivial direction.

Recent evidence for Mexico shows that conditional cash transfers can generate effects that go beyond the program's beneficiaries. Cash transfers to poor households can indirectly affect non poor households living in the same locality, where the transfers benefit the local economy at large. Angeluci and De Giorgi (2009) explore the effect of PROGRESA and found that non poor households in treated villages increase their consumption and benefit from the transfers by receiving more gifts and loans and by reducing their savings. A growing body of literature also emphasizes the role of social interactions in shaping outcomes. There is increasing evidence of the influence of neighborhoods and peers on outcomes such as child health and behavior (Case and Katz, 1991; Katz *et al.*, 2001), student outcomes (Sacerdote 2001, Lalive and Cattaneo, 2009), technology adoption in agriculture (Foster and Rosenzweig, 1995; Munshi, 2004) and retirement plan decisions (Duflo and Saez, 2002).

This paper explores the spillover effect of "Familias en Acción" on grade completion rates for ineligible households living in rural areas. The identification strategy exploits the fact that exposure to the program varies by locality of residence and year of birth. Estimates are based on difference-in-differences estimators that control for additive systematic variation both across localities and cohorts. When using cross section data, the selection of outcomes to be compared between cohorts needs to be not age dependent. For instance, it would be invalid to compare dropout rates between young and old cohorts as the dropout rate is higher for the older cohorts. Hence in this paper we will only consider completed years of education.

The sign of the net effect of the spillovers could be positive or negative. Positives effects can occur when, in the absence of formal credit markets, poor households share the grants with non poor households or when the transfers affect the local economy (Maluccio & Flores 2005, Coady & Harris 2001). Positive effects can also arise from imitation, peer pressure or from increased competition if poor children attain more years of education. The expected return to education of non poor children might change and they might decide to invest more in education to compensate the future increased competition. Negative effects can arise from congestion in schools because the program encourages poor children to remain in the school. My estimates measure the net causal effect of the program among non poor children.

The program "Familias en Acción" is targeted to families in the lowest level of the welfare index (level 1). However, a significant manipulation of the official score has been extensively reported in the literature (Camacho and Conover, 2009) where families that should have been classified in the level 2 are classified as level 1 in order to be eligible for the grant. To take this manipulation into account this study explores the spillover effect on families scoring on levels 3 to 6 of the reconstructed welfare index.

A growing body of the literature has documented the role of peers in individual schooling choices (Hanushek *et al.* 2003; Sacerdote 2001; Zimmerman 2003; Angrist and Lang 2004; Kremer, Miguel and Thornton 2004). As discussed in Manski (1993) and Moffit (2001), it is challenging to identify peer effects. Effects can be grouped into two main categories, which Manski (1993) defined as contextual and endogenous effects. On the one hand, these external effects may stem from some exogenous characteristics of the population. For example, average parental education in a given class may affect individual outcomes because children do homework together and benefit from the help received from their classmate's more educated parents when studying at their household. Alternatively, external effects may arise from individual outcomes. For example, if some students do better, they may help their classmates or they free up more teacher time to be devoted to more needy students. Another possibility is that doing well or poorly in school may become a social norm of a given population, to which members conform.

The second part of the paper uses the partial-population experiment setting created by the intervention in order to explore the role of peer effect (endogenous effect) on schooling decisions. In the spirit of Moffit (2001), identification of peer effects is achieved by studying how the treatment status of a specific subgroup of individuals -poor children eligible for the cash transfer - affects the outcomes of untreated individuals in the same group - non poor children in the same municipality and birth cohort. Given that "Familias en Acción" increases peer group school attendance while leaving unaffected the ineligible child's monetary incentive to attend school, this means that the response among ineligible children provides information on how strongly the peer group affects the individual. Identification requires that the program should not have had a direct effect on schooling decision of non poor children, imposing an arguable strong assumption. An implication of this assumption is that any effect of the program on the local economy will not affect directly the schooling outcomes of the non-poor and any effect will come through changes in the average outcome of the peers.

The closest work to the present paper is that of Lalive and Cattaneo (2009) and Bobonis & Finan (2009), who examine the effect of a conditional cash transfer program in Mexico (PROGRESA) on dropout and completed years of education for primary and secondary school, respectively. Both find that a 10 percentage point increase in peer participation raises own participation rates by 5 percentage points. Kremer, Miguel, and Thornton (2009) identify peer effects by taking advantage of experimental variation of an experiment that provides financial incentive to girls to perform well on exams. They found large spillover effects on the boys in the same schools despite the boys not being eligible for the cash rewards. Treated girls' scores rose on average by 0.29 standard deviations while (the ineligible) boys' scores rose by 0.16 standard deviations. This study will contribute to this literature exploring spillover effects on ineligible individuals of a program that incentivize school attendance of poor children.

The paper is organized as follows. Section 2.2 provides a brief discussion of the "Familias en Acción" program and its structure. Section 2.3 describes the data used in the analysis and section 2.4 discusses the identification strategy. Section 2.5 reports the main results for spillover effects, section 2.6 estimates endogenous social interactions, and section 2.7 presents falsification experiments. Section 2.8 concludes.

2.2 Program description

The program "Familias en Acción" (FA) is a large scale welfare program run by the Colombian government. It started in 2001 with the aim of fostering human capital accumulation. The program has three main components: education, nutrition, and health. The program gives a monetary transfer to mothers provided their children attend school regularly and are up to date with vaccinations, and growth and development monitoring visits. The monthly grant is set at 14,000 pesos (US$6) for each child attending primary school (grades 1-5), and 28,000 pesos (US$12) for each child attending secondary school (grades 6-11). Families with children below the age of 6 are entitled to receive a basic nutritional subsidy (approximately US$15 per month).

The targeting of the program took place in two stages. First, a set of municipalities were selected for the program in a non random fashion. Second, the poorest households in targeted areas with individuals aged 0–17 were eligible for the program.

Qualifying municipalities. A total of 622 out of the 1,098 municipalities in Colombia were deemed eligible to qualify for the program, on the basis of fulfilling the following criteria: (i) the municipality has sufficient education and health infrastructure; (ii) it has a bank; (iii) it has fewer than 100,000 inhabitants and is not a departmental capital, (iv) its administrative office has relatively up-to-date welfare lists and other important official documents.

Eligible households. Within each qualifying town, poor households with children aged 0–17 were eligible for the program Familias en Acción (FA). In the first two years of the program a total of 340,000 households were registered to participate (Attanasio et. al., 2006).

Colombia uses a proxy mean index, called SISBEN (Sistema de Identificación de Beneficiarios), to assign each household to one of six levels of the welfare indicator. The SISBEN index is the result of an algorithm that weighs households' variables associated with their socio-economic well-being. It consists of fourteen components measuring different aspects of household well-being (such as housing material, access to public utilities, ownership of durable assets, demographic composition, educational attainment, and labor force participation. On each dimension, households are classified according to mutually exclusive and collectively exhaustive categories with varying weights assigned to each category; these weights vary between urban and rural areas. For each household, the SISBEN score is then calculated by adding points across components. Possible scores range from 0 to 100 (with 0 being the most impoverished). This index is used to identify the most vulnerable population. In the case of FA, the program is available to families that score below 18 in rural areas and below 36 in urban areas which is considered the lowest welfare level (SISBEN level 1).

2.3 Data

The 2005 Census, conducted in Colombia between May 2005 and February 2006 by the National Administrative Department of Statistics (DANE), covers 96.3% of the total population. The data used in this paper consist of the 10% sample of dwellings provided by IPUMS International database at the University of Minnesota. The census sample includes 1,054,812 households and 4,117,607 individuals.

In the IPUMS sample covers only municipalities with a population of at least 20,000 persons in 1993 and assigns them a unique geographic identifier. Out of 1122 municipalities and other communities in Colombia, 900 did not exceed the minimum population. Consequently they were aggregated in 310 geographic units (between 2 to 5 communities per geographic unit). Those geographic units included qualifying municipalities and no eligible municipalities. In some cases a geographic unit combined municipalities where the program was and was not operating. Hence, my study will focus on the 222 municipalities that were not aggregated receiving a unique geographic

ID.

The final sample captures 119 municipalities where the program was operating (treatment municipalities) and 26 qualifying municipalities where the program was not operating (control municipalities). Control municipalities are communities where the program was not operating between 2001 and 2005 and had “reasonable” similar population size. Most of the control municipalities were towns with basic school and health infrastructure but without a bank or municipalities where the local authority did not register for the program.

2.4 Methodology

This section describes the identification strategy to estimate the spillover effects on ineligible children and to extend the analysis to measure peer effects on schooling outcomes.

2.4.1 Sources of variation

The program was targeted to poor families with children under 18. The date of birth and the municipality of birth jointly determine an individual’s exposure to the FA program. All children born between 1984 and 1995 (aged 6 to 17 in 2001)[25] in municipalities where the program was available were “exposed” to it. Poor children in treatment municipalities received a cash incentive to attend to the school. Non poor children did not received a grant but the availability of the program for a subgroup of the population -poor children- might affect their schooling outcomes as mentioned in the introduction (section 2.1).

The cohort born between 1984-1995 as young enough to be “exposed” to the program as they can be enrolled in primary or secondary education. We compare “exposed” cohorts with those born between 1971- 1982 (19-30 years old) that were too old to be “exposed” to the program.[26]

In order to split the population in poor (eligible) and non poor (non eligible), I replicate the poverty score described in section 2.2using the variables reported in the census. Given that the census does not include all the variables required to construct the index, I predict unavailable components using estimates from the Living Condition

[25]Children 0-5 years old in 2001 are excluded from the analysis as they were aged 4-9 at the time of the census and too young to be able to complete 3 or more years of education

[26]Individuals born in 1983 are excluded from the older cohort because they might be attending to secondary school.

Survey collected (1997 and 2003) and the Demographic and Health Survey 2005. [27]

Once SISBEN score is computed, I classify each household into one of the six welfare levels. Officially, the program was available to households classified in SISBEN level 1 but the manipulation of the score has been extensively reported in the literature (Camacho and Conover, 2009) where families that should have been classified as level 2 are classified as level 1. Hence, to take the manipulation into account, I consider that poor families are those classified in the SISBEN level 1 and 2. Therefore, I consider non poor (ineligibles) those households with a welfare score equivalent to level 3 to 6.

The relevant welfare index to compare between young and old cohorts is the one the individual had when was at school age. Consequently, that information is no longer available if the individual is not longer living with his/her parents. Hence, I exclude from the sample all individuals that reported to be the household head. The final sample in rural areas includes 41,715 individuals in schooling age and 21,527 in older cohorts.

Table 19 reports grade completion by age group and eligibility status (poor and non poor) . Birth cohorts are classified into four groups: (i) secondary school transition age (aged 6-11 in 2001), (ii) secondary school post-transition age (aged 12-14 in 2001), (iii) secondary school age (aged 15-17) and (iv) post-secondary school age (aged 19-24 and 25-30). The first column of each age group indicates the grade completion rate for municipalities where the program was not available between 2001-2005 (C) and each second column reports those for municipalities where the program was available (T). For cohorts that are not in school age - cannot benefit from the program - completion rates are slightly higher in control areas. [28]

2.4.2 Identification of spillover effects

The date of birth and the municipality of birth jointly determine the exposure to the program. To estimate the effect of the program among non poor (ineligible) children I estimate the following model

[27] Out of 14 categories required to compute SISBEN score, four components were incomplete or missing in the 2005 census: quality of roof material, family income, access to social protection and time to get water in rural areas. The first two components were predicted using the Living Condition Survey 1997, the third using the Living Condition Survey 2003 and the last one using the Demographic and Health Survey 2005. Quality level of roof material is predicted using estimated coefficients from ordered probit models that consider quality of floor material, number of rooms and regional dummy variables. Family income is predicted using household head's gender, age and level of education and regional dummy variables. Access to social protection considers type of work, age, gender, level of education and regional variables. Time to get water is predicted using water supply, floor material and regional variables.

[28] Completed primary school is equivalent to attain 5 or more years of education.

TABLE 19

Grade completion rates - 2005

	Age in 2001									
	6-11		12-14		15-17		19-24		25-30	
	C	T	C	T	C	T	C	T	C	T
A. Poor										
3 or more years of education	0.640	0.704	0.785	0.762	0.713	0.718	0.702	0.653	0.632	0.571
4 or more years of education	0.485	0.569	0.692	0.672	0.630	0.622	0.581	0.529	0.482	0.433
5 or more years of education	0.331	0.414	0.573	0.585	0.534	0.532	0.492	0.436	0.391	0.338
6 or more years of education	0.169	0.226	0.376	0.373	0.323	0.307	0.267	0.218	0.144	0.132
7 or more years of education	0.094	0.128	0.305	0.310	0.263	0.262	0.233	0.188	0.113	0.107
8 or more years of education	0.040	0.064	0.217	0.246	0.211	0.222	0.202	0.161	0.092	0.086
9 or more years of education	0.014	0.026	0.155	0.182	0.176	0.184	0.162	0.139	0.076	0.073
N	2154	24150	857	9354	635	7330	888	11008	568	7661
B. Non Poor										
3 or more years of education	0.913	0.921	0.957	0.956	0.949	0.944	0.929	0.915	0.901	0.890
4 or more years of education	0.825	0.838	0.938	0.928	0.932	0.907	0.892	0.861	0.839	0.809
5 or more years of education	0.640	0.693	0.914	0.896	0.913	0.867	0.843	0.805	0.774	0.735
6 or more years of education	0.461	0.485	0.839	0.752	0.784	0.690	0.688	0.603	0.536	0.468
7 or more years of education	0.306	0.324	0.775	0.686	0.729	0.636	0.630	0.555	0.481	0.416
8 or more years of education	0.170	0.189	0.699	0.603	0.669	0.588	0.577	0.509	0.427	0.367
9 or more years of education	0.076	0.084	0.590	0.507	0.621	0.532	0.534	0.470	0.380	0.326
N	2200	10510	974	4619	770	3542	1236	5474	880	3954

$$y_{icm} = \alpha_0 + \alpha_1 T_m * young_c + \alpha_2 X_{icm} + \eta_c + \lambda_m + \varepsilon_{icm} \quad (4)$$

where y measures grade completion for individual i in the birth cohort c and living in municipality m. Grade completion is a dummy variable that takes value 1 if the individual has completed the grade under consideration or an upper grade. T_m is a dummy variable that takes value 1 if the individual lives in a treatment municipality[29] and, $young_c$ takes value 1 if the individual was in school age (aged 6-17) when the program was launched. Hence, $T_m * young_c$ is equal to 1 if the individual was living in a municipality where the program was available while he/she was in school age. X_{icm} is a set of individual and family characteristics. η_c is a set of year of birth (cohort) fixed effect that controls for aggregate shocks that could affect y even in the absence of the program; λ_m is a set of municipality-of-birth fixed effect that account for any cohort-invariant unobservables within the municipality, and ε_{icm} is the error term.

The coefficient of interest is α_1, that can be expressed as the difference-in-differences estimator

$$\alpha_1 = E(y|T_m = 1, young = 1) - E(y|T_m = 0, young = 1) - E(y|T_m = 1, young = 0) + E(y|T_m = 0, young = 0)$$

[29]Migration between municipalities is low. In theory, non poor could migrate to municipalities where the program is not operating to reduce congestion in schools. To maintain exogeneity and reduce bias in the program effect, I consider the municipality of birth instead of current municipality as it will not be correlated with the program given than all individuals were born before the program started.

The OLS estimate of α_1 will consistently estimate the average effect of the program (spillover) among non poor individuals residing in treatment municipalities if the difference in unobservables between non poor cohorts are orthogonal to the availability of the program in the municipality, $E(\varepsilon_{icm}^{young} - \varepsilon_{icm}^{old} | T_m = 1) = E(\varepsilon_{icm}^{young} - \varepsilon_{icm}^{old} | T_m = 0)$.

Several reasons that might explain the existence of spillover effects on non poor children. First, grants to poor families can improve the local economy at large increasing income and access to formal/informal credit markets for non poor households. Second, noon poor individuals might imitate or react to the increased competition from poor individuals that stay longer in the school. Third, the program can create congestion in the school if poor children return/remain in the school increasing the class size. In addition, measures to mitigate the expected congestion in the school due to the increase in schooling demand of poor household might attract children from ineligible households.

2.4.3 Identification of social interactions (peer effect)

Spillover effects might be the result of different sources. Following Manski (1993), I can formalize the different channels through which the program can affect schooling outcomes on non poor children. In the linear in means model, the schooling outcome of child i in the peer group g is expressed as

$$y_{ig} = \alpha + \theta X_{ig} + \beta \bar{X}_g + \gamma \bar{Y}_g + u_{ig}$$

where β and γ capture the social interactions. Using Manski's terminology, β is the contextual effect that captures the role of peer characteristics on individual schooling outcomes, while γ is the endogenous peer effect and capture the effect of average group behavior on individual behavior. Finally, θ measures correlated effects where individuals in the same group tends to behave similarly because they face the same constraints or have similar characteristics.

As in Manski (1993), OLS estimation of the linear-in-means model cannot separately identify contextual and exogenous interactions. Policy interventions that create a partial population experiment, affecting only one part of the group, will allow separated identification of social interactions. Identification requires orthogonality between the policy variable the unobservables.

This section discusses the identification of social interactions based on a partial population experiment (Moffitt, 2001). Assuming that children schooling decisions follow the linear-in-means model, the outcome for poor and non poor can be expressed as

$$y^P_{icm} = \alpha + \theta X^P_{icm} + \beta \bar{X}_{cm} + \gamma \bar{Y}_{cm} + \delta T_m * young_c + \eta_c + \lambda_m + \varepsilon^P_{icm} \tag{5}$$

$$y^{NP}_{icm} = \alpha + \theta X^{NP}_{icm} + \beta \bar{X}_{cm} + \gamma \bar{Y}_{cm} + \eta_c + \lambda_m + \varepsilon^{NP}_{icm} \tag{6}$$

where y^P_{icm} and y^{NP}_{icm} are the schooling outcome (grade completion) of individual i in birth cohort c living in municipality m, for poor (P) and non-poor (NP) groups, respectively. X_{icm} and $\bar{X}_{cm}$ are individual and group exogenous characteristics, $\bar{Y}_{cm}$ is the average outcome of the reference group (excluding individual i) and $T_m * young_c$ is interaction between treatment locality and exposed cohort. $T_m * young_c$ is assumed not to have a direct effect on non-poor agents. ε_{icm} is an error term assumed to be independent and identically distributed. Group unobservables are assumed to be a linear combination of municipality (λ_m) and birth cohort unobservables (η_c). Then combining equations 5 and 6, and taking the group average

$$E(y_{icm}) = \alpha + \theta E(X_{icm}) + \beta \bar{X}_{cm} + \gamma \bar{Y}_{cm} + \delta T_m * young_c * E(P_{icm}) + \lambda_m + \eta_c \tag{7}$$

where $E(P_{icm})$ is the proportion of poor children in the birth cohort/municipality. If $\gamma \neq 1$, equation 7 has a unique solution:

$$\bar{Y}_{cm} = \frac{\alpha}{1-\gamma} + \frac{\theta+\beta}{1-\gamma}\bar{X}_{cm} + \frac{\delta}{1-\gamma}E(P_{icm}) * T_m * young_c + \lambda_m + \eta_c \tag{8}$$

Inserting equation 9 in equation 6, I obtain the reduced-form model

$$y^{NP}_{icm} = \frac{\alpha}{(1-\gamma)} + \theta X^{NP}_{icm} + \frac{\gamma\theta+\beta}{1-\gamma}\bar{X}_{cm} + \frac{\gamma\delta}{1-\gamma}E(P_{icm}) * T_m * young_c + \tilde{\lambda}_m + \tilde{\eta}_c + \varepsilon^{NP}_{icm} \tag{9}$$

In equation 9, $\frac{\gamma\delta}{1-\gamma}$ captures the effect of the program on non poor from changes in the outcome of poor (behavioral interactions or endogenous peer effect). Estimation of the effect on non-poor is based on difference in difference (DD) techniques. The DD-estimate is an unbiased estimate of the policy if, absent the policy change, the average-change between exposed (young) an unexposed (old) cohorts would have been the same for treatment and controls. Hence, the required identification assumption is that there are no omitted cohort-varying and municipality-specific effects correlated with the program. The difference between young and old cohorts can be interpreted as the growth in the schooling completion rate. As in Duflo (2000) an implication of the identification assumption can be tested, individuals that were not in school age when the program started (old), should not be affected by the policy. The increase in completion rates between old cohorts should not differ systematically across municipalities. The two parameters of interest are the overall effect on non poor children, $\phi \equiv \frac{\gamma\delta}{1-\gamma}$, and the endogenous peer-effect or behavioral response, γ.

Equation 6 is the equation to estimate . Equation 8 is equivalent to a first stage where group outcome is instrumented using the interaction between treatment municipality, exposed cohort and share of poor in the cohort ($E(P_{icm}) * T_m * young_c$). Then endogenous peer-effect are obtained by taking the ratio of the two DD and can be expressed as follow

$$\gamma_{IV} = \frac{(E[Y^{NP}_{icm}\big|T_m = 1, J = 1] - E[Y^{NP}_{icm}\big|T_m = 1, J = 0]) - (E[Y^{NP}_{icm}\big|T_m = 0, J = 1] - E[Y^{NP}_{icm}|T_m = 0, J = 0])}{(E[\bar{Y}_{cm}\big|T_m = 1, J = 1] - E[\bar{Y}_{cm}\big|T_m = 1, J = 0]) - (E[\bar{Y}_{cm}\big|T_m = 0, J = 1] - E[\bar{Y}_{cm}\big|T_m = 0, J = 0])} \tag{10}$$

where J is a dummy variable that takes value 1 if the cohort has been exposed to the program (young). IV estimations of the parameter γ will be consistent if the instrument satisfies two conditions: (i) it is strongly correlated with $\bar{Y}$ and (ii) it is uncorrelated with Y^{NP} beyond the direct effect through $\bar{Y}$ (no correlation between the instrument and error term ε_{icm}). Condition (i) can be tested by the first stage (8) but condition (ii) cannot be directly tested and will be maintained as an assumption.

2.5 Results

In this section I explore the net effect of the transfers on grade completion rates for non-poor (ineligible) individuals and discuss the potential mechanisms involved. According to the impact evaluation for the program "Familias en Acción" (Attanasio *et al.* (2004) and Attanasio *et al.* (2010)) the impact of the program on enrollment was higher in rural areas and for children around the transition between primary and secondary. Those results are consistent with my estimates for completion rates in rural and urban areas. Hence, this section will focus on rural areas and the estimates for urban areas will be reported in the Appendix II . To facilitate comparisons, I start reporting the effect on poor individuals living in municipalities where the program was available. [30]

[30] Note that the group of poor individuals include households classified as sisben welfare level 1 and 2. As explained before, the program is eligible for families officially classified as level 1 in the SISBEN score but due to the severe manipulation, many households that should have been classified as level 2 were classified as level 1. To take into account the manipulation I consider them as poor.

TABLE 20

Difference in difference effects among poor

	Years of Education						
	3+	4+	5+	6+	7+	8+	9+
	(1)	(2)	(3)	(4)	(6)	(7)	(8)
Panel A							
Aged 6-11 in 2001 x Treatment	**0.0798*****	**0.0998*****	**0.1115*****	**0.0828*****	**0.0598****	**0.0508****	0.0278
	(0.0259)	(0.0342)	(0.0354)	(0.0306)	(0.0243)	(0.0249)	(0.0215)
Aged 12-14 in 2001 x Treatment	-0.0073	-0.0043	0.0403	0.0245	0.0299	0.0517	0.0390
	(0.0239)	(0.0289)	(0.0309)	(0.0400)	(0.0396)	(0.0339)	(0.0257)
Aged 15-17 in 2001 x Treatment	**0.0351***	0.0200	**0.0413***	0.0109	0.0260	**0.0349***	0.0169
	(0.0205)	(0.0225)	(0.0246)	(0.0271)	(0.0194)	(0.0207)	(0.0232)
Observations	58,398	58,398	58,398	58,383	58,383	58,383	58,383
R-squared	0.2708	0.2572	0.2511	0.2153	0.1963	0.1740	0.1517
Panel B							
Aged 6-8 in 2001 x Treatment	**0.1111*****	**0.1100*****	**0.0993****	**0.0666***	0.0356	0.0275	0.0172
	(0.0317)	(0.0366)	(0.0419)	(0.0373)	(0.0351)	(0.0317)	(0.0254)
Aged 9-11 in 2001 x Treatment	**0.0464***	**0.0889****	**0.1245*****	**0.1002****	**0.0857*****	**0.0758*****	**0.0391****
	(0.0237)	(0.0372)	(0.0445)	(0.0463)	(0.0288)	(0.0222)	(0.0188)
Aged 12-14 in 2001 x Treatment	-0.0073	-0.0043	0.0402	0.0244	0.0298	0.0517	0.0390
	(0.0239)	(0.0289)	(0.0309)	(0.0399)	(0.0396)	(0.0340)	(0.0257)
Aged 15-17 in 2001 x Treatment	**0.0352***	0.0200	**0.0413***	0.0108	0.0259	**0.0348***	0.0168
	(0.0205)	(0.0225)	(0.0246)	(0.0270)	(0.0194)	(0.0207)	(0.0232)
Observations	58,398	58,398	58,398	58,383	58,383	58,383	58,383
R-squared	0.2709	0.2573	0.2512	0.2153	0.1965	0.1742	0.1517

Note: Dependent variables are dummies for completion of a minimum number of years of education. Dependent variable in column (1) is a dummy variable for 3 or more years of education. All specifications include municipality of residence in 2000, year of birth dummies, gender, mother's education, welfare index, welfare index interacted with treatment municipality, share of poor children and share of males at cohort/municipality level. Robust standard errors are clustered at municipality level and reported in parenthesis. * Significant at 10%, ** significant at 5%, and *** significant at 1%.

Program's effect among the poor

Table 20 reports the intention-to-treat effect on poor children based on equation (4) but for poor children. Estimates will combine the direct effect of the policy and any indirect effect due to spillover effects. The probability of acquiring primary education or more years of education increases by 11.2 percentage points if children are in the transition age while I will show next that this effect is 9 percentage points for non poor children (table 21). In panel B, the estimates for the cohort around the transition is disaggregated in two groups, aged 6-8 and 9-11 in 2001 only to improve the comparison with the results on poor available in the literature. Attanasio (2004) finds the program has effect on children that were in secondary school age at the time of the follow up (2003) and this will be comparable with my cohort 9-11 in 2001. Table 20 (panel B) reports that individuals in the first group are likely to be in the transition between primary and secondary school at the time of the census (2005) so there is effect on completing primary school grades but not yet in upper years. For the second group (9-11 in 2001) that is in year 7-12 or has dropped out the school, I find the highest effects are observed on completion of primary school (12 percentage points) and 6 years of education (10 percentage points).

TABLE 21

Difference in differences estimates among non poor

	Years of Education						
	3+	4+	5+	6+	7+	8+	9+
	(1)	(2)	(3)	(4)	(6)	(7)	(8)
Aged 6-11 in 2001 x Treatment	0.0158	**0.0389***	**0.0897*****	**0.0992****	**0.0864****	**0.0839****	**0.0718***
	(0.0129)	(0.0210)	(0.0287)	(0.0388)	(0.0360)	(0.0385)	(0.0368)
Aged 12-14 in 2001 x Treatment	**0.0180***	**0.0318****	**0.0360***	0.0076	-0.0051	-0.0187	-0.0125
	(0.0097)	(0.0139)	(0.0192)	(0.0260)	(0.0265)	(0.0283)	(0.0288)
Aged 15-17 in 2001 x Treatment	0.0136	0.0177	0.0066	-0.0046	-0.0101	-0.0064	-0.0198
	(0.0115)	(0.0135)	(0.0166)	(0.0192)	(0.0172)	(0.0208)	(0.0254)
Observations	32,659	32,659	32,659	32,648	32,648	32,648	32,648
R-squared	0.0629	0.1210	0.2315	0.2846	0.3101	0.3172	0.3161

Note: Dependent variables are dummies for completion of a minimum number of years of education. Dependent variable in column (1) is a dummy variable for 3 or more years of education. All specifications include municipality of residence in 2000, year of birth dummies, gender, mother's education, welfare index, share of poor children and share of males at cohort/municipality level. Robust standard errors are clustered at municipality level and reported in parenthesis. * Significant at 10%, ** significant at 5%, and *** significant at 1%.

Spillover effects among non poor

This section reports the results on spillover effect among non poor after four years since the program started. Table 21 reports the OLS estimates for spillover effects (α_1) from equation (4) considering three groups of individuals "exposed" to the program. [31] In this context, exposure is an increasing function of individual´s date of birth. Table 21 shows the grant aiming to incentivize poor children to stay or return to the school has increased the education of non-poor children. Its effect is stronger among individuals exposed to spillovers during the transition between primary and secondary school (aged 6-11 in 2001), with effects in the range of 7 to 10 percentage points. Considering the probability of completing primary school (5 or more years of education), I find the youngest cohort experienced an increment by 9 percentage points, the cohort aged 12-14 in 2001 increases by 3.6 percentage points, whilst there is no effect on the cohort that was exposed long after the transition period between primary and secondary school. These results suggest that the program might also reduce the dropout rate for non poor individuals in the transition age between primary and secondary school and incentivize individuals between 12-14 to finish primary school.[32]

To further investigate this issue, I include an interaction between exposure to the program in treatment areas and share of poor individuals in the cohort/municipality. Increments in class size, and the consequent congestion, are more likely the higher the proportion of eligible individuals in the cohort/municipality is. In table 22, I explore

[31]To take into account possible correlations within each municipality, standard errors are clustered at the municipality level throughout the analysis.

[32]Given enrollment is age dependent and the identification strategy requires the comparison with cohort in post school age, it is not possible to compute estimates for the enrollment rates.

TABLE 22

Difference in difference effect among non poor II

	Years of Education						
	3+	4+	5+	6+	7+	8+	9+
	(1)	(2)	(3)	(4)	(6)	(7)	(8)
Aged 6-11 in 2001 x Treatment x proportion of poor	**0.0503*****	**0.0629****	**0.1203*****	**0.1429*****	**0.1272****	**0.1356*****	**0.1098****
	(0.0145)	(0.0249)	(0.0336)	(0.0477)	(0.0492)	(0.0519)	(0.0528)
Aged 12-14 in 2001 x Treatment x proportion of poor	0.0256	0.0234	0.0406	-0.0039	0.0253	0.0166	0.0268
	(0.0163)	(0.0223)	(0.0304)	(0.0395)	(0.0419)	(0.0448)	(0.0458)
Aged 15-17 in 2001 x Treatmentx proportion of poor	0.0156	0.0221	0.0454	0.0261	0.0107	0.0100	0.0167
	(0.0196)	(0.0245)	(0.0283)	(0.0368)	(0.0397)	(0.0400)	(0.0393)
Aged 6-11 in 2001 x Treatment	-0.0098	0.0071	0.0314	0.0318	0.0262	0.0197	0.0184
	(0.0138)	(0.0239)	(0.0329)	(0.0442)	(0.0426)	(0.0461)	(0.0445)
Aged 12-14 in 2001 x Treatment	0.0055	0.0205	0.0176	0.0116	-0.0153	-0.0247	-0.0245
	(0.0116)	(0.0170)	(0.0238)	(0.0327)	(0.0340)	(0.0357)	(0.0359)
Aged 15-17 in 2001 x Treatment	0.0071	0.0082	-0.0136	-0.0166	-0.0148	-0.0108	-0.0269
	(0.0141)	(0.0171)	(0.0209)	(0.0266)	(0.0260)	(0.0289)	(0.0314)
Observations	32,659	32,659	32,659	32,648	32,648	32,648	32,648
R-squared	0.0634	0.1214	0.2324	0.2857	0.3108	0.3181	0.3167

Note: Dependent variables are dummies for completion of a minimum number of years of education. Dependent variable in column (1) is a dummy variable for 3 or more years of education. All specifications include municipality of residence in 2000, year of birth dummies, gender, mother's education, welfare index, share of poor children and its interaction with treatment municipality, and share of males at cohort/municipality level. * Significant at 10%, ** significant at 5%, and *** significant at 1%.

the spillover effects allowing differential effects of the program based on the share of poor individuals. I find that increased fraction of poor children in treatment municipalities increases the completion rates of non poor children in school transition age by 5 to 14 percentage points. This increment on grade completion suggests that congestion was not severe enough to discourage non poor children to attend or that do not not play a significant role in schooling decision.

To control for cohort size, in table 23 I replicate estimates from tables 21 and 22. After controlling for birth year/municipality cohort size, the conclusions remain. [33]

Finding the net spillover is positive and an increasing function of the share of poor children at cohort-municipality level signals that congestion is not a dominating effect. Then there are two potential dominating effects. First, the effect of the program in the local economy. The transfer may affect the goods market through the increased expenditures from poor households increasing either the goods prices in the treatment municipalities or the sales from non-poor to the poor. In the risk-sharing model - in which agents fully insure against idiosyncratic risk by pooling resources and consuming a fixed share of total income - consumption is independent of their individual income, conditional on aggregate resources (Townsend, 1994). One of the implications of this model is that, given two individuals 1 and 2, an increase in agent 1's income will increase aggregate resources, resulting in higher consumption for both agents. This

[33]There is a slight increment in the coefficients for the interaction between treatment municipality, exposed cohort, and share of poor at cohort/municipality level.

TABLE 23

Difference in difference effects among non poor III

	Years of Education						
	3+	4+	5+	6+	7+	8+	9+
	(1)	(2)	(3)	(4)	(6)	(7)	(8)
A.							
Aged 6-11 in 2001 x Treatment	0.0133	0.0323	**0.0835*** **	**0.1021*** **	**0.0872** **	**0.0856** **	**0.0766** **
	(0.0130)	(0.0210)	(0.0288)	(0.0390)	(0.0364)	(0.0388)	(0.0367)
Aged 12-14 in 2001 x Treatment	**0.0166* **	**0.0281** **	**0.0325* **	0.0092	-0.0047	-0.0178	-0.0098
	(0.0097)	(0.0140)	(0.0195)	(0.0262)	(0.0268)	(0.0284)	(0.0288)
Aged 15-17 in 2001 x Treatment	0.0125	0.0148	0.0038	-0.0034	-0.0097	-0.0057	-0.0177
	(0.0114)	(0.0132)	(0.0166)	(0.0195)	(0.0175)	(0.0210)	(0.0255)
Cohort size	0.0001	0.0004	0.0003	-0.0002	-0.0000	-0.0001	-0.0003
	(0.0001)	(0.0003)	(0.0004)	(0.0003)	(0.0003)	(0.0004)	(0.0003)
Observations	32,659	32,659	32,659	32,648	32,648	32,648	32,648
R-squared	0.0630	0.1213	0.2317	0.2847	0.3101	0.3172	0.3161
B.							
Aged 6-11 in 2001 x Treatment x proportion of poor	**0.0509*** **	**0.0469* **	**0.1205*** **	**0.1921*** **	**0.1638*** **	**0.1786*** **	**0.1614*** **
	(0.0153)	(0.0265)	(0.0338)	(0.0456)	(0.0490)	(0.0508)	(0.0486)
Aged 12-14 in 2001 x Treatment x proportion of poor	0.0261	0.0133	0.0408	0.0270	0.0483	0.0436	0.0592
	(0.0163)	(0.0219)	(0.0300)	(0.0407)	(0.0439)	(0.0476)	(0.0477)
Aged 15-17 in 2001 x Treatmentx proportion of poor	0.0158	0.0161	0.0455	0.0446	0.0244	0.0262	0.0360
	(0.0195)	(0.0238)	(0.0287)	(0.0383)	(0.0412)	(0.0421)	(0.0411)
Aged 6-11 in 2001 x Treatment	-0.0099	0.0106	0.0313	0.0210	0.0182	0.0103	0.0071
	(0.0137)	(0.0239)	(0.0328)	(0.0436)	(0.0421)	(0.0455)	(0.0436)
Aged 12-14 in 2001 x Treatment	0.0054	0.0227	0.0175	0.0046	-0.0204	-0.0307	-0.0317
	(0.0116)	(0.0174)	(0.0238)	(0.0321)	(0.0339)	(0.0359)	(0.0360)
Aged 15-17 in 2001 x Treatment	0.0070	0.0090	-0.0137	-0.0191	-0.0167	-0.0131	-0.0296
	(0.0141)	(0.0171)	(0.0209)	(0.0269)	(0.0262)	(0.0291)	(0.0315)
Cohort size	-0.0000	0.0002	-0.0000	**-0.0007** **	**-0.0005* **	**-0.0006* **	**-0.0008*** **
	(0.0001)	(0.0003)	(0.0004)	(0.0003)	(0.0003)	(0.0003)	(0.0003)
Observations	32,659	32,659	32,659	32,648	32,648	32,648	32,648
R-squared	0.0634	0.1215	0.2324	0.2861	0.3110	0.3184	0.3171

Note: Panel A and B replicate the last two tables adding a control variable for the size of the birth year cohort in the municipality. Robust standard errors are clustered at municipality level and reported in parenthesis. * Significant at 10%, ** significant at 5%, and *** significant at 1%.

efficient resource allocation can be achieved through a series of informal loans and transfers (Fafchamps and Lund (2003) in Phillipines, Angeluci and De Giorgi (2009) in Mexico). Therefore, the higher income for agent 1 will also result in an increase in net transfers to agent 2.

A second dominating effect is endogenous *peer effect.* This effect could take place due to imitation, competition or peer pressure. There are strategic complementarities in peer participation and effort in the education, if the child enrolls in school, the time that peers spend in class, as well as their effort levels inside and outside the classroom, can enhance the child's learning, in addition to his or her own ability and the school environment. At the same time, there might be a desire to conform with others due to peer pressure or social norms, resulting in children not wanting to deviate from choices made by others in her reference group (quadratic conformist utility function).

2.6 Endogenous social interactions (*Peer effect*)

Due to data constraints, I am unable to explore separately the different mechanisms involved. In this section I exploit the partial population experiment introduced by the program to estimate the endogenous social interactions(*peer effect*) under the assumption that an increase in completion rates among non poor children in treatment municipalities is the result of the exogenous increase in completion rates among poor children within the municipality, not the result of changes in contextual or individual variables affected by the program. This is an strong assumption and can not be tested. Hence, the results should be interpreted with caution.

Estimates in this section are based on equation 6 and 8. As explained in section 2.4, partial population experiment creates an exogenous variation on group average completion rates. I instrument average grade completion rate (poor and non poor children) for each age group using the interaction between exposed cohort, treatment municipality and proportion of poor children at cohort/municipality level, and the interaction between the first two variables. Coefficients for the first stage and the F test for joint significance are reported in table 24.

Table 25 reports the IV estimates of endogenous social effects (γ) from equation 8. For simplicity, I present in the same row the coefficients of endogenous social effect for different outcomes. They are not directly comparable because they refer to different grades. A 10 percentage point increase in peer group grade completion rate leads to a 3.6 to 6.7 percentage point increase in individual grade completion for non poor. The highest peer group effects are observed for completing primary education of a highest grade (column 3) and for completing 6 or more years of education (column 4).

TABLE 24

First stage - non poor

	Group proportion - Years of Education						
	3+	4+	5+	6+	7+	8+	9+
	(1)	(2)	(3)	(4)	(6)	(7)	(8)
Aged 6-11 in 2001 x Treatment x proportion of poor	**0.1096*** **	**0.1516*** **	**0.2098*** **	**0.2265*** **	**0.2577*** **	**0.3055*** **	**0.3234*** **
	(0.0181)	(0.0286)	(0.0347)	(0.0373)	(0.0398)	(0.0406)	(0.0401)
Aged 12-14 in 2001 x Treatment x proportion of poor	**0.1091*** **	**0.1184*** **	**0.1052*** **	-0.0210	-0.0312	-0.0437	-0.0474
	(0.0209)	(0.0247)	(0.0261)	(0.0298)	(0.0306)	(0.0315)	(0.0344)
Aged 15-17 in 2001 x Treatmentx proportion of poor	**0.0845*** **	**0.0933*** **	**0.0773*** **	**-0.0545* **	**-0.0654** **	**-0.0824*** **	**-0.0934*** **
	(0.0201)	(0.0238)	(0.0266)	(0.0308)	(0.0311)	(0.0312)	(0.0266)
Aged 6-11 in 2001 x Treatment	-0.0092	-0.0121	0.0085	0.0077	-0.0118	-0.0283	-0.0397
	(0.0152)	(0.0259)	(0.0369)	(0.0449)	(0.0466)	(0.0498)	(0.0490)
Aged 12-14 in 2001 x Treatment	**-0.0278* **	-0.0206	-0.0140	0.0041	-0.0177	-0.0144	0.0002
	(0.0157)	(0.0213)	(0.0234)	(0.0262)	(0.0263)	(0.0252)	(0.0269)
Aged 15-17 in 2001 x Treatment	-0.0175	-0.0234	-0.0285	-0.0012	-0.0027	-0.0013	-0.0005
	(0.0156)	(0.0200)	(0.0228)	(0.0249)	(0.0263)	(0.0281)	(0.0311)
Observations	32,768	32,768	32,768	32,768	32,768	32,768	32,768
R-squared	0.6690	0.7253	0.7957	0.8081	0.8028	0.8038	0.7938
F - test	7.643	6.640	8.659	13.779	15.559	20.336	22.758
Prob > F	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Note: Dependent variables are the completion rate in the group of a minimum number of years of education. Dependent variable in column (1) is a continuous variable for the proportion of individuals in the cohort-municipality with 3 or more years of education. All specifications include municipality of residence in 2000, year of birth dummies, gender, mother's education, welfare index, share of poor children and its interaction with treatment municipality, and share of males at cohort/municipality level. Robust standard errors are clustered at municipality level and reported in parenthesis. * Significant at 10%, ** significant at 5%, and *** significant at 1%.

TABLE 25

Endogenous social interaction effects (IV estimates - non poor)

	Years of Education						
	3+	4+	5+	6+	7+	8+	9+
	(1)	(2)	(3)	(4)	(6)	(7)	(8)
Group mean	**0.3657*** **	**0.4487*** **	**0.6765*** **	**0.6438*** **	**0.5063*** **	**0.4567*** **	**0.3506*** **
	(0.1137)	(0.1076)	(0.0741)	(0.0738)	(0.0782)	(0.0777)	(0.0908)
Observations	32,659	32,659	32,659	32,648	32,648	32,648	32,648
R-squared	0.0808	0.1452	0.2569	0.3126	0.3374	0.3436	0.3380

Note: Group mean is the proportion of children (poor and non-poor) in the cohort/municipality that has at least the years of education specified in the title of the column. Dependent variables are dummies for completion of a minimum number of years of education. Dependent variable in column (1) is a dummy variable for 3 or more years of education. All specifications include municipality of residence in 2000, year of birth dummies, gender, mother's education, welfare index, share of poor children and its interaction with treatment municipality, and share of males at cohort/municipality level. Robust standard errors are clustered at municipality level and reported in parenthesis. * Significant at 10%, ** significant at 5%, and *** significant at 1%.

TABLE 26
Control experiment among non poor

	Years of Education						
	3+	4+	5+	6+	7+	8+	9+
	(1)	(2)	(3)	(4)	(6)	(7)	(8)
Aged 19-24 in 2001 x Treatment	0.0035	0.0120	0.0181	-0.0014	0.0049	0.0076	0.0040
	(0.0144)	(0.0190)	(0.0249)	(0.0212)	(0.0225)	(0.0202)	(0.0191)
Observations	10,990	10,990	10,990	10,985	10,985	10,985	10,985
R-squared	0.0596	0.0905	0.1268	0.2258	0.2265	0.2383	0.2395

Note: Considers individuals between 25-30 years old in 2001 as the base category. Dependent variables are dummies for completion of a minimum number of years of education. Dependent variable in column (1) is a dummy variable for 3 or more years of education. All specifications include municipality of residence in 2000, year of birth dummies, gender, mother's education, welfare index, share of poor children and share of males at cohort/municipality level. Robust standard errors are clustered at municipality level and reported in parenthesis. * Significant at 10%, ** significant at 5%, and *** significant at 1%.

2.7 Robustness check

Control experiment

The identification strategy to estimate spillover effects relies on the assumption that in the absence of the program, the increment in completion rates from old to young cohorts would not have been systematically different in treatment and control municipalities. Given that individuals aged 19 or older in 2001 were not exposed to the availability of the program during the schooling age, an implication of the identification assumption can be tested. The increase in completion rates between cohort groups should not differ systematically across municipalities. Table 26 shows a falsification experiment where individuals between 19-24 are assumed to be exposed to the program and they are compared against individuals between 25-30 years in 2001. The sample excludes individuals that report to be the household head. The estimated differences in differences are very close to 0.

2.8 Conclusions

In 2001, Colombia launched a conditional cash transfer program aiming to increase human capital among poor households. The program was targeted geographically and allocated to poor households according to a poverty score (proxy mean index). The subsidy provided incentives to poor student to remain in the school.

This paper explores the indirect effects of cash grants allocated to poor children on grade completion rates among non-poor households in rural areas. On average, the estimates indicate the program leads to an increase in primary completion rate among non-poor children by 9 percentage points. For basic secondary education (6 to 9 years), the causal effect of the intervention on grade completion rates is between 7 to 10 percentage points for cohorts exposed during the transition between primary and secondary school.

Data limitations restrict the analysis on the mechanisms involved but I can rule out congestion as the dominating effect. Using the partial population experiment setting created by the intervention and assuming no direct effect of the grants on non poor children, I estimate the role of endogenous social interactions on schooling decisions. A 10 percentage point increase in peer group grade completion rate leads to a 3.6 to 6.7 percentage point increase in individual grade completion for non poor. These results suggest that the endogenous social interactions might play an important role in schooling decisions and that policy intervention might have social multiplier effects.

Appendix II - Results on urban areas

This appendix replicates the results for urban areas where the program had smaller effects.

TABLE 27

Grade completion rate - Urban

	Age in 2001									
	6-11		12-14		15-17		19-24		25-30	
	C	T	C	T	C	T	C	T	C	T
A. Poor										
3 or more	0.814	0.863	0.904	0.914	0.890	0.894	0.843	0.846	0.779	0.792
4 or more	0.683	0.746	0.860	0.878	0.828	0.847	0.776	0.780	0.687	0.703
5 or more	0.510	0.586	0.802	0.828	0.766	0.792	0.710	0.722	0.611	0.633
6 or more	0.338	0.400	0.675	0.711	0.612	0.659	0.528	0.552	0.403	0.414
7 or more	0.200	0.244	0.586	0.633	0.531	0.599	0.470	0.497	0.345	0.353
8 or more	0.105	0.131	0.476	0.530	0.478	0.535	0.418	0.443	0.288	0.301
9 or more	0.040	0.055	0.347	0.418	0.413	0.470	0.364	0.399	0.243	0.262
N	3233	17936	1353	7331	951	5468	1398	8116	881	5574
B. Non Poor										
3 or more	0.968	0.969	0.988	0.986	0.989	0.984	0.981	0.977	0.969	0.968
4 or more	0.902	0.916	0.979	0.982	0.982	0.977	0.969	0.966	0.950	0.945
5 or more	0.765	0.790	0.968	0.973	0.971	0.967	0.957	0.953	0.923	0.924
6 or more	0.581	0.618	0.932	0.937	0.931	0.925	0.896	0.885	0.798	0.814
7 or more	0.393	0.437	0.889	0.902	0.891	0.899	0.869	0.852	0.749	0.775
8 or more	0.236	0.275	0.829	0.846	0.858	0.861	0.835	0.818	0.707	0.735
9 or more	0.113	0.135	0.740	0.755	0.817	0.824	0.802	0.781	0.666	0.695
N	5392	22325	2440	9823	1959	7469	3212	12180	2270	9207

TABLE 28

Difference in difference effects among poor

	Years of Education						
	3+	4+	5+	6+	7+	8+	9+
	(1)	(2)	(3)	(4)	(6)	(7)	(8)
Panel A							
Aged 6-11 in 2001 x Treatment	**0.0381****	**0.0512****	**0.0551***	**0.0455**	**0.0270**	**0.0048**	-0.0137
	(0.0190)	(0.0224)	(0.0284)	(0.0327)	(0.0251)	(0.0245)	(0.0228)
Aged 12-14 in 2001 x Treatment	-0.0045	-0.0025	-0.0027	0.0018	0.0154	0.0194	0.0335
	(0.0170)	(0.0200)	(0.0214)	(0.0262)	(0.0282)	(0.0302)	(0.0292)
Aged 15-17 in 2001 x Treatment	**-0.0072**	0.0025	**-0.0027**	0.0169	0.0392**	**0.0274**	0.0184
	(0.0136)	(0.0164)	(0.0173)	(0.0257)	(0.0195)	(0.0218)	(0.0256)
Observations	48,118	48,118	48,118	48,111	48,111	48,111	48,111
R-squared	0.1063	0.1675	0.2486	0.2807	0.2939	0.2861	0.2696
Panel B							
Aged 6-8 in 2001 x Treatment	**0.0670*****	**0.0703*****	**0.0580****	**0.0252**	0.0008	-0.0209	-0.0311
	(0.0242)	(0.0249)	(0.0290)	(0.0328)	(0.0245)	(0.0249)	(0.0228)
Aged 9-11 in 2001 x Treatment	**0.0085**	**0.0317**	**0.0522***	**0.0662***	**0.0537***	**0.0312**	**0.0040**
	(0.0156)	(0.0223)	(0.0310)	(0.0358)	(0.0294)	(0.0271)	(0.0242)
Aged 12-14 in 2001 x Treatment	-0.0045	-0.0025	-0.0027	0.0019	0.0155	0.0195	0.0336
	(0.0170)	(0.0200)	(0.0214)	(0.0262)	(0.0282)	(0.0302)	(0.0292)
Aged 15-17 in 2001 x Treatment	-0.0072	0.0025	-0.0027	0.0169	**0.0393****	0.0274	0.0185
	(0.0136)	(0.0164)	(0.0173)	(0.0257)	(0.0196)	(0.0218)	(0.0256)
Observations	48,118	48,118	48,118	48,111	48,111	48,111	48,111
R-squared	0.1067	0.1676	0.2486	0.2808	0.2941	0.2863	0.2696

Note: Dependent variables are dummies for completion of a minimum number of years of education. Dependent variable in column (1) is a dummy variable for 3 or more years of education. All specifications include municipality of residence in 2000, year of birth dummies, gender, mother's education, welfare index, welfare index interacted with treatment municipality, share of poor children and share of males at cohort/municipality level. Robust standard errors are clustered at municipality level and reported in parenthesis. * Significant at 10%, ** significant at 5%, and *** significant at 1%.

TABLE 29

Difference in difference effects among non poor

	Years of Education						
	3+	4+	5+	6+	7+	8+	9+
	(1)	(2)	(3)	(4)	(6)	(7)	(8)
Aged 6-11 in 2001 x Treatment	0.0015	**0.0144***	**0.0255****	**0.0345****	**0.0379****	**0.0323***	0.0188
	(0.0038)	(0.0083)	(0.0121)	(0.0172)	(0.0166)	(0.0179)	(0.0205)
Aged 12-14 in 2001 x Treatment	-0.0015	0.0038	0.0063	0.0063	0.0136	0.0190	0.0172
	(0.0034)	(0.0045)	(0.0059)	(0.0113)	(0.0127)	(0.0124)	(0.0138)
Aged 15-17 in 2001 x Treatment	-0.0031	-0.0012	-0.0016	-0.0045	0.0062	0.0016	0.0067
	(0.0040)	(0.0045)	(0.0051)	(0.0094)	(0.0097)	(0.0115)	(0.0131)
Observations	73,813	73,813	73,813	73,801	73,801	73,801	73,801
R-squared	0.0277	0.1167	0.3283	0.3974	0.4622	0.4847	0.4880

Note: Dependent variables are dummies for completion of a minimum number of years of education. Dependent variable in column (1) is a dummy variable for 3 or more years of education. All specifications include municipality of residence in 2000, year of birth dummies, gender, mother's education, welfare index, share of poor children and share of males at cohort/municipality level. Robust standard errors are clustered at municipality level and reported in parenthesis. * Significant at 10%, ** significant at 5%, and *** significant at 1%.

TABLE 30

Difference in difference effect among non poor II

	Years of Education						
	3+	4+	5+	6+	7+	8+	9+
	(1)	(2)	(3)	(4)	(6)	(7)	(8)
Aged 6-11 in 2001 x Treatment x proportion of poor	**0.0174****	**0.0726*****	**0.1423*****	**0.1253*****	**0.1260*****	**0.0904****	0.0446
	(0.0081)	(0.0206)	(0.0309)	(0.0380)	(0.0391)	(0.0408)	(0.0427)
Aged 12-14 in 2001 x Treatment x proportion of poor	**0.0163***	0.0136	-0.0100	**-0.0553****	**-0.0588****	**-0.0597****	**-0.0637***
	(0.0095)	(0.0128)	(0.0150)	(0.0228)	(0.0252)	(0.0277)	(0.0348)
Aged 15-17 in 2001 x Treatmentx proportion of poor	0.0128	**0.0188***	0.0185	-0.0029	-0.0048	-0.0226	-0.0421
	(0.0083)	(0.0107)	(0.0136)	(0.0200)	(0.0254)	(0.0296)	(0.0272)
Aged 6-11 in 2001 x Treatment	-0.0055	-0.0121	-0.0199	-0.0067	-0.0047	0.0044	0.0043
	(0.0048)	(0.0118)	(0.0173)	(0.0230)	(0.0232)	(0.0245)	(0.0274)
Aged 12-14 in 2001 x Treatment	**-0.0077***	-0.0019	0.0129	**0.0266***	**0.0342****	**0.0418*****	**0.0393****
	(0.0047)	(0.0067)	(0.0087)	(0.0142)	(0.0165)	(0.0156)	(0.0190)
Aged 15-17 in 2001 x Treatment	-0.0075	-0.0076	-0.0054	-0.0023	0.0086	0.0107	0.0207
	(0.0047)	(0.0056)	(0.0069)	(0.0111)	(0.0131)	(0.0148)	(0.0155)
Observations	73,813	73,813	73,813	73,801	73,801	73,801	73,801
R-squared	0.0278	0.1173	0.3298	0.3982	0.4629	0.4851	0.4882

Note: Dependent variables are dummies for completion of a minimum number of years of education. Dependent variable in column (1) is a dummy variable for 3 or more years of education. All specifications include municipality of residence in 2000, year of birth dummies, gender, mother's education, welfare index, share of poor children and its interaction with treatment municipality, and share of males at cohort/municipality level. * Significant at 10%, ** significant at 5%, and *** significant at 1%.

TABLE 31

Difference in difference effect among non poor III

	Years of Education						
	3+	4+	5+	6+	7+	8+	9+
	(1)	(2)	(3)	(4)	(6)	(7)	(8)
A.							
Aged 6-11 in 2001 x Treatment	0.0015	**0.0144***	**0.0250****	**0.0342****	**0.0372****	**0.0321***	0.0177
	(0.0038)	(0.0082)	(0.0120)	(0.0171)	(0.0164)	(0.0179)	(0.0201)
Aged 12-14 in 2001 x Treatment	-0.0015	0.0038	0.0061	0.0062	0.0135	0.0189	0.0169
	(0.0034)	(0.0045)	(0.0059)	(0.0113)	(0.0128)	(0.0125)	(0.0137)
Aged 15-17 in 2001 x Treatment	-0.0031	-0.0011	-0.0014	-0.0044	0.0065	0.0017	0.0072
	(0.0040)	(0.0045)	(0.0050)	(0.0094)	(0.0097)	(0.0114)	(0.0130)
Cohort size	-0.0000	0.0000	0.0002	0.0001	0.0002	0.0001	0.0004
	(0.0001)	(0.0002)	(0.0002)	(0.0003)	(0.0002)	(0.0002)	(0.0002)
Observations	73,813	73,813	73,813	73,801	73,801	73,801	73,801
R-squared	0.0277	0.1167	0.3284	0.3974	0.4622	0.4847	0.4880
B.							
Aged 6-11 in 2001 x Treatment x proportion of poor	**0.0175****	**0.0726*****	**0.1418*****	**0.1249*****	**0.1251*****	**0.0903****	0.0429
	(0.0082)	(0.0208)	(0.0311)	(0.0382)	(0.0391)	(0.0408)	(0.0426)
Aged 12-14 in 2001 x Treatment x proportion of poor	**0.0163***	0.0136	-0.0103	**-0.0556****	**-0.0595****	**-0.0598****	**-0.0650***
	(0.0095)	(0.0128)	(0.0149)	(0.0229)	(0.0251)	(0.0277)	(0.0346)
Aged 15-17 in 2001 x Treatmentx proportion of poor	0.0127	**0.0188***	0.0190	-0.0026	-0.0041	-0.0225	-0.0405
	(0.0083)	(0.0108)	(0.0137)	(0.0198)	(0.0252)	(0.0296)	(0.0273)
Aged 6-11 in 2001 x Treatment	-0.0055	-0.0121	-0.0200	-0.0068	-0.0050	0.0044	0.0038
	(0.0048)	(0.0118)	(0.0173)	(0.0230)	(0.0232)	(0.0245)	(0.0273)
Aged 12-14 in 2001 x Treatment	**-0.0077**	-0.0019	0.0129	**0.0266***	**0.0342****	**0.0418*****	**0.0393****
	(0.0047)	(0.0067)	(0.0087)	(0.0142)	(0.0165)	(0.0156)	(0.0190)
Aged 15-17 in 2001 x Treatment	-0.0075	-0.0076	-0.0054	-0.0023	0.0085	0.0107	0.0206
	(0.0047)	(0.0056)	(0.0069)	(0.0111)	(0.0131)	(0.0148)	(0.0155)
Cohort size	-0.0000	-0.0000	0.0001	0.0001	0.0002	0.0000	0.0004
	(0.0001)	(0.0001)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Observations	73,813	73,813	73,813	73,801	73,801	73,801	73,801
R-squared	0.0278	0.1173	0.3298	0.3982	0.4629	0.4851	0.4882

Note: Panel A and B replicate the last two tables adding a control variable for the size of the birth year cohort in the municipality. Robust standard errors are clustered at municipality level and reported in parenthesis. * Significant at 10%, ** significant at 5%, and *** significant at 1%.

TABLE 32

First stage - non poor

	Group proportion - Years of Education						
	3+	4+	5+	6+	7+	8+	9+
	(1)	(2)	(3)	(4)	(6)	(7)	(8)
Aged 6-11 in 2001 x Treatment x proportion of poor	**0.0655*** **	**0.1166*** **	**0.1800*** **	**0.2465*** **	**0.2974*** **	**0.3315*** **	**0.3562*** **
	(0.0127)	(0.0253)	(0.0331)	(0.0415)	(0.0419)	(0.0423)	(0.0444)
Aged 12-14 in 2001 x Treatment x proportion of poor	**0.1019*** **	**0.1304*** **	**0.1348*** **	**0.1070*** **	**0.0879*** **	**0.0437* **	0.0362
	(0.0106)	(0.0141)	(0.0162)	(0.0240)	(0.0260)	(0.0249)	(0.0324)
Aged 15-17 in 2001 x Treatmentx proportion of poor	**0.0650*** **	**0.0894*** **	**0.1053*** **	**0.0884*** **	**0.0654*** **	0.0399	0.0129
	(0.0117)	(0.0146)	(0.0159)	(0.0204)	(0.0234)	(0.0258)	(0.0239)
Aged 6-11 in 2001 x Treatment	-0.0071	-0.0110	-0.0125	-0.0273	-0.0439	-0.0513	-0.0684*
	(0.0056)	(0.0129)	(0.0187)	(0.0276)	(0.0282)	(0.0316)	(0.0360)
Aged 12-14 in 2001 x Treatment	**-0.0251*** **	**-0.0270*** **	**-0.0164* **	-0.0008	0.0080	**0.0295* **	0.0322
	(0.0072)	(0.0094)	(0.0097)	(0.0162)	(0.0169)	(0.0163)	(0.0213)
Aged 15-17 in 2001 x Treatment	**-0.0168*** **	**-0.0195** **	**-0.0201** **	-0.0174	0.0003	0.0059	0.0148
	(0.0062)	(0.0084)	(0.0091)	(0.0130)	(0.0126)	(0.0139)	(0.0148)
Observations	73,850	73,850	73,850	73,850	73,850	73,850	73,850
R-squared	0.5731	0.7349	0.8736	0.8927	0.9176	0.9206	0.9215
F - test	16.446	15.341	14.873	8.223	10.458	14.389	17.855
Prob > F	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Note: Dependent variables are the completion rate in the group of a minimum number of years of education. Dependent variable in column (1) is a continuous variable for the proportion of individuals in the cohort-municipality with 3 or more years of education. All specifications include municipality of residence in 2000, year of birth dummies, gender, mother's education, welfare index, share of poor children and its interaction with treatment municipality, and share of males at cohort/municipality level. Robust standard errors are clustered at municipality level and reported in parenthesis. * Significant at 10%, ** significant at 5%, and *** significant at 1%.

TABLE 33

IV estimates - non poor

	Years of Education						
	3+	4+	5+	6+	7+	8+	9+
	(1)	(2)	(3)	(4)	(6)	(7)	(8)
Group mean	**0.1794** **	**0.4226*** **	**0.6756*** **	**0.5790*** **	**0.5126*** **	**0.3703*** **	**0.2090*** **
	(0.0783)	(0.0746)	(0.0599)	(0.0651)	(0.0623)	(0.0689)	(0.0785)
Observations	73,813	73,813	73,813	73,801	73,801	73,801	73,801
R-squared	0.0346	0.1404	0.3552	0.4188	0.4782	0.4970	0.4956

Note: Group mean is the proportion of children (poor and non-poor) in the cohort/municipality that has at least the years of education specified in the title of the column. Dependent variables are dummies for completion of a minimum number of years of education. Dependent variable in column (1) is a dummy variable for 3 or more years of education. All specifications include municipality of residence in 2000, year of birth dummies, gender, mother's education, welfare index, share of poor children and its interaction with treatment municipality, and share of males at cohort/municipality level. Robust standard errors are clustered at municipality level and reported in parenthesis. * Significant at 10%, ** significant at 5%, and *** significant at 1%.

TABLE 34

Control Experiment on non poor

	Years of Education						
	3+	4+	5+	6+	7+	8+	9+
	(1)	(2)	(3)	(4)	(6)	(7)	(8)
Aged 19-24 in 2001 x Treatment	-0.0015	0.0039	-0.0024	**-0.0207***	**-0.0385*****	**-0.0403*****	**-0.0457*****
	(0.0058)	(0.0068)	(0.0075)	(0.0119)	(0.0129)	(0.0129)	(0.0140)
Observations	25,917	25,917	25,917	25,913	25,913	25,913	25,913
R-squared	0.0233	0.0389	0.0532	0.1189	0.1434	0.1630	0.1829

Note: Considers individuals between 25-30 years old in 2001 as the base category. Dependent variables are dummies for completion of a minimum number of years of education. Dependent variable in column (1) is a dummy variable for 3 or more years of education. All specifications include municipality of residence in 2000, year of birth dummies, gender, mother's education, welfare index, share of poor children and share of males at cohort/municipality level. Robust standard errors are clustered at municipality level and reported in parenthesis. * Significant at 10%, ** significant at 5%, and *** significant at 1%.

3 Housing Upgrade and Health: A Randomized Experiment in Community Nurseries

Abstract

In this paper we use a randomized experiment to study the impact on child health and adult life satisfaction of a program improving the physical infrastructure of nurseries in the city of Cartagena, Colombia. We find that replacing floors, upgrading toilets, kitchen, and play areas has no impact on child health but these results are subject to a high level of non-selective attrition on children. We also find that the program improves caregiver's mental health as measured by the CES-D depression score. Size effects are moderate on average, at around 0.3 to 0.4 standard deviations, with stronger effects for those between the 30 and 70 percentile of the depression score distribution.

3.1 Introduction

Housing conditions affects health in many different ways. Deficient housing conditions on aspects such as water, sanitation, and safe food preparation and storage, can be a critical factor behind the rapid spread of communicable and food-borne diseases. In addition, housing and neighborhoods can be also considered as a psychosocial environment that can affect mental health and life satisfaction. These are some of the reasons why countries may devote substantial resources to upgrade slum areas and improve housing quality, particularly for the poorest groups in the population. However, despite the importance of housing as a factor influencing well-being, limited work has been done to assess the causal impact of housing improvement programs on health and welfare.

In this paper we use a randomized experiment to examine the causal effect of improving housing conditions on child health and adult mental health. The experiment was carried-out in the city of Cartagena, Colombia, and the intervention consisted in upgrading floors, walls, and toilets. The program was specifically targeted to houses accommodating a state-funded community nursery called *"Hogares Comunitarios"*.

Our paper contributes to expand the still scant literature on the physical and mental benefits of housing improvements. Contributions in this area come from both the medical and the economic literature. The medical literature has documented some association between housing, and physical and mental health. Thomson *et al.* (2009) conducts a systematic review in this area and reports that housing improvements, especially warmth improvements, can produce health benefits among adults and children;

they also report some positive effects on mental health. In turn, a still limited body of the economics literature has been focused on the identification of the impact of housing upgrade on physical and mental health. Katz, Kling, and Liebman (2001) examine the impact of changes in residential neighborhoods on the well-being of families in high-poverty areas that received housing vouchers through a random lottery in the United States. They find that households that were offered vouchers experienced improvements in multiple measures of well-being, including health among heads of household, and a reduction in the likelihood of asthma attacks and injuries among children. Devoto *et al.* (2012) find that facilitating access to credit for private tap leads to a positive impact on mental health and perceived quality of life but has not impact on income or health outcomes. Cattaneo *et al.* (2009) investigate the impact of a large-scale Mexican program to replace dirt floors with cement floors on child health and adult happiness. They find that children living in improved houses experience lower incidence of parasitic infestations, diarrhea, and anemia, as well as an improvement in cognitive development; they also find an improvement in adult welfare as measured by depression scores, perceived stress scales, and self-reported satisfaction with their housing and quality of life.

Anticipating the results that we discuss in the following sections, we find improving housing conditions has no significant effect in the available outcomes of child health. Several factor may be conditioning this result. To begin with, sample drop-out rates reached almost 61% of children initially attending community nurseries. Although this attrition appears to be randomly distributed among children in both treatment and control groups, it still nonetheless reduces the power of our statistics for this particular group (children). Moreover, data availability restrictions prevents us from using direct measures of the health outcomes of interest (such as parasitic infestations). Instead we use indirect outcomes (diarrhea, respiratory diseases and other illnesses), what dampens the power of our identification strategy for this particular group. Finally, the data was collected nine months after the intervention, which may be too short of a time-frame to capture health effects.

In addition, we find that the intervention under consideration leads to a statistically significant improvement in mental health measures for adults living in houses subjected to the treatment. Mental health is described by measures of the current level of depressive symptomatology including depressed mood, feelings of helplessness and hopelessness, psycho-motor retardation, loss of appetite, and sleep disturbance (Center for Epidemiologic Studies Depression Scale).

Finally, a word of caution regarding the validity of our results. Although the internal validity of the results is guaranteed by the experimental design of the evaluation (that is, the findings are valid in the context of the evaluation), they do not necessarily apply or can be extrapolated to other cities with different socioeconomic characteristics, climatic conditions, and/or sanitary infrastructure.

The rest of the paper is organized as follows. Section 3.2 describes the intervention, while section 3.3 discusses the data used and presents basic statistics. Section 3.4 explains the framework to estimate the impact of the housing upgrades on child health and caregiver depression. The main results are presented in section 3.5, and section 3.6 discusses the robustness of our results. Finally, section 3.7 concludes.

3.2 Background

Hogares Comunitarios (HC) is a large program of community nurseries introduced all over Colombia in the mid 1980s. The program is run by the *Instituto Colombiano de Bienestar Familiar* (ICBF). Each of these nurseries provide nutrition and childcare to an average of 13 poor children between 0-6 years old. They operate in the household of the caregiver (*Madre Comunitaria).* Currently, there are more than 80,000 nurseries across the country.

In 2007, the ICBF launched a pilot program to upgrade 2,700 *Hogares Comunitarios*. The program improved the quality of floors in the kitchen, toilets, common area and playgrounds, replacing rough wood, board, cement and dirt floors with tile. Additionally, the program improved the quality of the walls and roof and built new toilets. In 2008, the ICBF agreed to randomly select what nurseries would be improved in the city of Cartagena were the improvements were not taking place.

Upgrades on community nurseries in Cartagena, Colombia's fifth largest city, took place between December 2008 and March 2009. The initial plan for the evaluation considered six areas. Unfortunately, for three of these six areas the improvements began before the baseline survey could be carried-out. As a consequence of this, we had to exclude those three areas from our study. For the remaining three areas, nurseries were randomly allocated to treatment and control groups. Of the 150 nurseries deemed to require physical improvements by ICBF engineers, 72 were randomized-in and 78 were randomized-out. The control nurseries received the intervention in 2010.

3.3 Data

3.3.1 Data characteristics and sources

The ICBF collected data on nurseries and children on November 2008 before the housing improvement had taken place, and in November 2009 roughly nine months after the improvements finished. It collected data on 150 nurseries. The nursery module included in the survey gathered specific information on the caregiver, including socioeconomic characteristics, specific training, knowledge on diarrhea and other illnesses, and depressive symptoms. The module also includes very detailed information on the nursery's (household) access to public services (water, sewage, electricity, etc), and infrastructure, including the quality of ceilings, walls and floors, number of rooms and toilets, and floor material composition for the kitchen, dinning, toilets and playground.

The child and family module includes data socioeconomic characteristics, access to social programs, child health outcomes (diarrhea, respiratory diseases, and other illnesses in the 7 days prior the interview), health services utilization, history of attendance to *Hogares Comunitarios* nurseries, and distance to important places in the town -such as the nearest health center and school.

Most of the data collection was carried out in the nursery. In very few cases data was collected at the child household. As we explain below in more detail, the follow up survey suffered from a high level of attrition of children in the initial sample.

3.3.2 Outcome measures

We are interested in capturing the effects of this housing improvements on both the mental health of the caregiver, and specific outcomes of child health. As mention before, we do not observe in the data the direct outcome of interest (such as parasitic infestations) but rather indirect ones (diarrhea, respiratory diseases and other illnesses).

Child Health. Child health is maternal-reported, and includes the prevalence of diarrhea, respiratory disease and other illnesses in the 7 days prior to the survey.

Mental Health Adult mental health is measured using the Center for Epidemiologic Studies Depression Scale (CES-D Scale) designed by Radloff (1977) to measure depression. The CES-D is a 20-item self-report instrument that assesses the presence and severity of depressive symptoms occurring over the past week. Questions are easy to answer and cover most of the areas included in the diagnostic criteria for depression, reflecting depressed mood, feelings of guilt and worthlessness, feelings of helplessness and hopelessness, psychomotor retardation, loss of appetite, and sleep disturbance. Response categories indicate the frequency of occurrence of each item, and are scored on a 4-point scale: 0 = rarely or none of the time, 1 = some or a little of the time, 2 = occasionally or a moderate amount of the time, and 3 = most or all of the time. Four of the items were positively worded while the others were negatively worded. We obtained the CES-D score by reversing the scores of the answers to the positively worded items and then summing up the scores across the answers for the 20 items. Therefore, a higher score denotes a higher level of perceived depression measured on a scale of 0 to 60. In general, scores of 16 and above are indicative of high depressive symptoms and 21 and above are indicative of severe clinical depression. It is important to note that in this paper we will focus on the depression scores instead of the depression status.

3.3.3 Balance between treatment and control group

In 2008, we collected data on 150 nurseries. A year after that, we could collect data on 149 of those nurseries.[34] In the follow up survey, only 145 caregivers answered the depression module. We restricted our sample to those nurseries with no missing data on depression in both interviews, obtaining a final sample comprising 140 nurseries. Table 35 compares the balance in household and caregiver characteristics for nurseries in the treatment and control groups. It also compares pretreatment outcomes on depression. None of the differences are statistically significant at a 5%.[35]

Table 36 shows the difference in means for all the children attending to 149 nurseries in baseline. Balance between groups is crucial for the validity of the randomized experiment. At a 5% significance level, we detect imbalances between treatment and control groups in only 2 out of 19 variables.

3.3.4 Attrition

As we mentioned earlier, we observed a sizable attrition for children in the sample, but we do not find statistical evidence indicating a differential selection process for the treatment and control groups. In other words, the process driving the attrition appears to be orthogonal to the intervention. ICBF officials consider that this high level of attrition is due to the creation of a new preschool for children between 3-5 years old.

The quantitative characteristics of this process are as follows. Out of 1,958 children interviewed in 149 nurseries at the baseline survey, 865 stayed in the nursery at the

[34]One nursery in the control group was temporary closed due to the medical condition on the caregiver that needed a surgery.

[35]We find similar results for the difference in means at baseline using the sample of 150 nurseries.

TABLE 35

Difference of means for nurseries in 2008

	Mean		N	Dif	St. Error	t- stat.
	T	C				
Housing characteristics						
ICBF infrastructure score	101.8	102.0	140	-0.8	3.6	-0.2
Quality wall (=1 if high quality materials)	1.0	0.9	140	0.0	0.0	0.7
Quality roof (=1 if high quality materials)	0.2	0.1	138	0.0	0.1	0.8
Number of toilets	1.1	1.1	140	0.1	0.1	1.0
Toilet with conection to sewage	0.8	0.7	140	0.1	0.1	1.0
Toilet outside the household	0.3	0.4	139	-0.1	0.1	-1.5
Exclusive toilet for the nursery	0.1	0.1	139	0.0	0.1	0.3
Floors with low quality at:						
Kitchen	0.5	0.6	139	0.0	0.1	-0.5
Dinning room	0.6	0.7	139	-0.1	0.1	-1.1
Playground	0.6	0.7	139	-0.1	0.1	-1.4
Main toilet	0.3	0.4	140	-0.1	0.1	-1.5
Other household characteristics						
Poverty index (SISBEN)	0.9	1.0	135	-0.1	0.0	-1.8
Distance to center (minutes)	47.4	45.6	140	1.7	2.9	0.6
Distance to pharmacy (minutes)	10.8	12.3	140	-1.4	1.2	-1.2
Distance to school (minutes)	9.4	10.3	140	-0.6	1.2	-0.5
Distance to hospital (minutes)	31.3	33.3	140	-1.2	3.3	-0.4
Distance to health center (minutes)	13.7	13.9	139	-0.1	1.5	-0.1
Nursery caregiver characteristics						
Age	43.4	45.6	140	-2.3	1.6	-1.4
Years of education	10.2	9.7	139	0.5	0.5	0.9
Years of experience as caregiver	14.0	14.6	140	-0.7	1.2	-0.6
Nursery caregiver pre-treatment outcomes						
Depression score	6.4	6.0	140	0.5	1.0	0.5
Depression (=1 if had depression)	0.1	0.1	140	0.0	0.0	0.1

TABLE 36

Difference of means in 2008. Children

	Mean T	Mean C	N	Dif	St. Errors	t-stat
Family and household characteristics						
Tutor's Age	29.37	29.82	1,866	-0.43	(0.423)	-1.025
Tutor's Education	8.06	7.93	1,861	0.12	(0.243)	0.477
Head's Age	34.26	34.55	1,804	-0.26	(0.667)	-0.388
Head's Education	7.67	7.75	1,809	-0.11	(0.252)	-0.436
Poor (SISBEN 1)	0.93	0.96	1,624	-0.04	(0.020)	**-2.000**
Time to centre	46.93	47.20	1,870	-0.18	(2.264)	-0.081
Time to pharmacy	12.49	12.67	1,870	-0.14	(0.870)	-0.161
Time to school	11.46	10.81	1,870	0.83	(0.810)	1.021
Time to hospital	35.78	36.59	1,870	-0.17	(2.871)	-0.057
Child characteristics						
Male	0.52	0.53	1,958	-0.01	(0.024)	-0.216
Age in months	42.34	42.84	1,860	-0.34	(0.637)	-0.540
Health Service utilization						
Preventive care utilization, last 12 months	0.83	0.85	1,870	-0.02	(0.024)	-1.014
Preventive care utilization, last 30 days	0.30	0.25	1,566	0.05	(0.031)	1.486
Up to date on health controls	0.89	0.84	1,552	0.05	(0.022)	**2.120**
Health insurance	0.91	0.92	1,863	-0.02	(0.017)	-1.000
Pre-treatment health outcomes						
Prevalence of diarrhea	0.10	0.12	1,870	-0.02	(0.018)	-1.332
Prevalence of respiratory diseases	0.50	0.53	1,870	-0.03	(0.032)	-0.846
Prevalence of other illness	0.09	0.13	1,869	-0.04	(0.021)	-1.962
Prevalence of any illness	0.56	0.61	1,870	-0.05	(0.031)	-1.467

TABLE 37

Attrition

	Attrition Rates T	Attrition Rates C	Dif.	St. Error	t-stat
(1)	0.547	0.569	-0.020	0.026	-0.785
(2)	0.601	0.620	-0.017	0.025	-0.649

Note: (1) reports attrition rate from children that were not longer enrolled in the nursery at the follow up survey. (2) reports the attrition rate from (1) and also includes children that were enrolled but absent on the day of the follow up interview. Sample size: 946 children in treatment group and 1012 in control group.

follow up and 1,093 were not longer enrolled. Table 37 reports the attrition between baseline and follow up. Row (1) indicates the attrition rate due to no enrollment one year after the baseline. The attrition was 55% in the treatment group and 57% in the control group but the difference was not statistically significant. In addition, there is a second source related to the absence on the day of the follow up interview. Around 11.5% of the 865 that stayed in the nursery at the follow up didn't report symptoms for the illnesses covered in the survey.[36] Row (2) takes into account the attrition due to absence on the day of the interview. Now the attrition rates increases to 60% for the treatment group and 62% to the control group. Thus, as already said, after comparing nurseries in the treatment and control groups we do not find significant differences between them.

We further investigate if there is evidence of a differential selection processes underlying the observed attrition in both the control and treatment groups (something that may bias the results). Table 38 reports the differences in means for children that were not enrolled or absent on the day of the follow up interview. We observe that the sample is balanced across intervention and control areas, with imbalances detected on 2 out to 19 variables. These results are consistent with the assumption of a common selection process, orthogonal to the intervention, affecting nursery attendance equally in the two group. In other words, we find no evidence of a selective attrition.

Finally, it is important to note that, although the attrition appears not to be selective (children that dropped out have similar characteristics irrespective of being in the control or treatment groups), we do observe some unbalance between those children that stayed in the treatment and control nurseries. This is a direct consequence of the reduction in sample size that followed the attrition. Had the sample remained closer to the initial size, this difference would have a much smaller weight in the results. Table 39 reports the difference in means at baseline for those children that were interviewed in baseline and follow up survey. Children in control nurseries have a higher prevalence of diarrhea, are less likely to be up to date on health check up and those differences are significant at the 5% level. We also find children in control nurseries experiencing more illnesses, excluding diarrhea and respiratory diseases, but the difference is only significant at 10%. We will discuss this results in section 3.5.

[36] According to the nursery caregiver, 95% of the 103 children that didn't report illnesses attended to the nursery in the last 5 days.

TABLE 38

Difference of means at baseline for children not interviewed in 2009

	Mean		N	Dif	St. Errors	t-stat
	T	C				
Family and household characteristics						
Tutor's Age	29.22	30.36	1,115	-1.14	(0.527)	**-2.167**
Tutor's Education	7.99	7.83	1,112	0.13	(0.296)	0.441
Head's Age	34.67	35.01	1,076	-0.33	(0.857)	-0.388
Head's Education	7.72	7.70	1,080	-0.02	(0.283)	-0.055
Poor (SISBEN 1)	0.91	0.96	957	-0.06	(0.025)	**-2.211**
Time to centre	46.10	46.62	1,119	-0.54	(2.333)	-0.230
Time to pharmacy	12.32	12.46	1,119	-0.07	(0.849)	-0.088
Time to school	11.61	10.64	1,119	1.14	(0.845)	1.353
Time to hospital	35.48	35.91	1,119	0.21	(2.815)	0.075
Child characteristics						
Male	0.51	0.51	1,196	0.00	(0.028)	-0.006
Age in months	46.41	46.60	1,114	-0.12	(0.874)	-0.141
Health Service utilization						
visits to health services - 12m	0.83	0.84	1,119	-0.01	(0.028)	-0.447
visits to health services - 30d	0.31	0.26	931	0.05	(0.034)	1.360
Up to date - health check up	0.88	0.84	919	0.04	(0.026)	1.670
Health insurance	0.90	0.92	1,113	-0.03	(0.022)	-1.291
Pre-treatment health outcomes						
Prevalence of diarrhea	0.10	0.11	1,119	0.00	(0.023)	-0.169
Prevalence of respiratory diseases	0.49	0.51	1,119	-0.01	(0.036)	-0.387
Prevalence of other illness	0.10	0.13	1,119	-0.03	(0.025)	-1.258
Prevalence of any illness	0.55	0.58	1,119	-0.03	(0.037)	-0.854

TABLE 39

Difference in means at baseline for children interviewed in 2009

	Mean		N	Dif	St. Errors	t-stat
	T	C				
Family and household characteristics						
Tutor's Age	29.24	29.01	848	0.28	(0.581)	0.485
Tutor's Education	8.02	7.98	846	0.05	(0.301)	0.151
Head's Age	33.76	34.12	820	-0.25	(0.872)	-0.286
Head's Education	7.51	7.74	822	-0.28	(0.324)	-0.876
Poor (SISBEN 1)	0.94	0.97	746	-0.03	(0.020)	-1.296
Time to centre	48.09	48.02	849	0.22	(2.406)	0.091
Time to pharmacy	12.77	13.00	849	-0.23	(1.090)	-0.208
Time to school	11.52	10.87	849	0.80	(0.947)	0.849
Time to hospital	36.44	36.88	849	-0.11	(3.150)	-0.036
Child characteristics						
Male	0.53	0.54	865	-0.01	(0.038)	-0.177
Age in months	36.82	37.17	843	-0.24	(0.750)	-0.325
Health Service utilization						
Preventive care utilization, last 12 months	0.82	0.86	849	-0.05	(0.032)	-1.584
Preventive care utilization, last 30 days	0.30	0.24	714	0.06	(0.042)	1.372
Up to date on health controls	0.90	0.83	711	0.07	(0.030)	**2.272**
Health insurance	0.91	0.92	848	-0.01	(0.021)	-0.389
Pre-treatment health outcomes						
Prevalence of diarrhea	0.09	0.14	849	-0.05	(0.025)	**-2.154**
Prevalence of respiratory diseases	0.52	0.57	849	-0.05	(0.042)	-1.083
Prevalence of other illness	0.08	0.13	848	-0.04	(0.027)	-1.671
Prevalence of any illness	0.58	0.65	849	-0.06	(0.038)	-1.679

3.4 Empirical Framework

Effect on children

Although the identification of the treatment effect relies on the randomization, in the case of children we consider double differences instead of simple difference because children in the control group seem to be more prompt to diseases (see table 39). In section 3.3.3 and 3.6, we will show there is a high level of drop out but no evidence of a selective attrition. We used a fixed effect approach that controls for preexisting differences between children in treatment and control groups

$$Y_{ihst} = \alpha + \beta T_h + \delta T_h * after_t + \phi after_t + \gamma X_{iht} + \lambda_s + \varepsilon_{ihst} \tag{11}$$

where Y_{iht} is the health outcome for child i at time t in nursery h located in stratum s. T_h is a dummy variable which equals 1 if the nursery was selected for the intervention. $after$ is a dummy variable that indicates if the data belongs to the follow up survey. X is a vector of child and family variables such as age in months, age-squared, gender, mother's age and years of education (lineal and squared). λ_s is a set of dummy variables for the stratum.ε_{ihst} is an error term which is uncorrelated with the error term of children attending to other nurseries, but which may be correlated with other children in the same nursery.

In this context, β captures a combination of differences from attrition and any other preexisting differences between treatment and control group (*group effect*). ϕ captures aggregate shocks that affect the outcome variable in the same way for both treatment and control groups (*time effect*).δ is the parameter of interest and captures the causal effect of the intervention on child health.

Effect on the caregiver

The randomized experiment provides us with a credible source of identification to estimate the effects of the intervention. For the case of the caregiver, we estimate the following model

$$Y_{sm} = \alpha + \beta T_{sm} + \gamma X_{sm} + \lambda_s + \varepsilon_m \tag{12}$$

where Y_{sm} is the depression score for the caregiver m in the nursery s one year after the intervention. [37] T is a dummy variable which equals 1 if the caregiver's

[37] Alternative transformation of the depression score has been considered: logarithms, box cox and inverse hyperbolic transformation.

household (nursery) was selected for the intervention. X is a vector of individual-level variables such as age, age-squared, years of education and years of education-squared.[38] λ_s is a set of dummy variables for the stratified randomization.$\varepsilon_m$is an error term. Randomization guarantees that conditional on the stratum the error term is uncorrelated with the treatment status, hence $E(\varepsilon_m|\lambda_s, T = 1) = E(\varepsilon_m|\lambda_s, T = 0)$.

We start estimating equation 12 by OLS. However, the distribution of the mental health score may change in a way that is not revealed by an examination of averages. Therefore, a parameter of interest in the presence of heterogeneous treatment effects is the quantile treatment effect (QTE). As originally defined by Doksum (1974) and Lehmann (1974), the QTE corresponds, for any fixed percentile, to the horizontal distance between two cumulative distribution functions. This result assumes rank preservation requiring that the relative value (rank) of the potential outcome for a given individual to be the same regardless of whether that individual is in the treatment or in the control groups.[39]

The estimation of quantile treatment effects (QTEs) allows us to analyze the effects on the entire distribution. We distinguish between conditional and unconditional effects. Under exogenous treatment, conditional quantile treatment effects (CQTEs) are defined conditional on the value of the set of covariates X, whereas unconditional effects summarize the causal effect of a treatment for the entire population. Unconditional QTEs name is used to distinguish them from commonly used conditional quantile regressions (Koenker and Bassett (1978), Koenker (2005)). The "unconditional quantiles" are the quantiles of the marginal distribution of the outcome variable Y. Our estimates will be based on the UQTEs as described below and presented in section 3.5. Estimations of conditional QTEs will be discussed in Appendix III.

Let Y^0 and Y^1 the potential outcomes with and without treatment and $F(Y^0)$ and $F(Y^1)$ the corresponding distributions. Y_i is the observed outcome, which is $Y_i \equiv Y_i^1 T_i + Y_i^0(1 - T_i)$.

Conditional Quantile Treatment Effect We start with the standard model for linear quantile regression. We assume that Y is a linear function in X and T,

$$Y_i = X_i\beta^\tau + T\delta^\tau + \varepsilon_i \text{ and } Q^\tau_{\varepsilon_i} = 0$$

for $i = 1, ..., n$ and $T\varepsilon(0, 1)$. $Q^\tau_{\varepsilon_i}$ refers to the τth quantile of the unobserved random variable ε_i.

[38] Notice that the randomness of the program allocation guarantees the consistency of the estimates, and the variables in X_{sm} are included to increase precision and not to remove bias.

[39] Note this problem doesn't occur when estimating the average treatment effects because difference in means coincides with the means of the differences.

In addition, we assume that conditional on X, T and (Y_0, Y_1) are independent ($Y_0, Y_1 \perp T|X$). Both assumptions imply

$$Q^{\tau}_{Y|X,T} = X_i\beta^{\tau} + T\delta^{\tau}$$

where $Q^{\tau}_{Y|X} \equiv F^{-1}_{Y|X}(\tau)$. δ^{τ} represents the CQTEs at quantile τ. δ^{τ} and β^{τ} coefficients can thus be estimated by the classical quantile regression estimator suggested by Koenker and Bassett (1978).

Unconditional Quantile Treatment Effect Following Firpo (2007), the unconditional QTE for the quantile τ is given by $\Delta^{\tau} = Q^{\tau}_{Y_1} - Q^{\tau}_{Y_0}$. In this case, quantile treatment effects are simple differences between quantiles of the marginal distributions of potential outcomes. However, if rank preservation holds, then the simple difference in quantiles turn out to be the quantiles of the treatment effect.

The two identification assumptions are that:

- selection to treatment is based on observable variables (exogeneity assumption), $Y_0, Y_1 \perp T|X$
- common support of X, $0 < Pr[T = 1|X = x] < 1$, that is both treatment assignment levels have a positive probability of occurrence.

An additional assumption is for some values of $\tau\varepsilon(0,1)$, quantiles are well defined and unique.

Firpo (2007) proposes a two step approach that is a reweighed version of the procedure proposed by Koenker and Bassett (1978) for quantile estimation. This estimation technique requires a nonparametric first step in which the propensity score is estimated. The final estimators will be equal to the differences between two quantiles, which can be expressed as solutions of minimization problems, where the minimands are sums of check functions, which are convex empirical processes.

The weighting estimator for Δ^{τ} therefore is

$$(\alpha, \Delta^{\tau}) = \underset{\alpha,\Delta^{\tau}}{argmin} \sum W^F_i \times \rho_{\tau}(Y_i - \alpha - T_i\Delta)$$

$$W^F_i = \frac{T_i}{Pr(T = 1|Xi)} + \frac{1 - T_i}{1 - Pr(T = 1|Xi)}$$

One of the advantages of using UQTE over CQTE is that the definition of UQTE does not change when we change the set of covariates X. These covariates are included in the first-step regression and then integrated out. However, the definition of the effects is not a function of the covariates.

3.5 Results

3.5.1 Effect on child health

Table 40 presents the effect of housing upgrades on child health using a fixed effect approach. As we described in section 3.3.3 the level of attrition was high but presumably not selective. However the results should be interpreted carefully due to the potential unbalance at baseline on health outcomes presented in table 39 and the lack of power. The program seems to have no impact on any of the child health outcomes (diarrhea, respiratory diseases, other illnesses and any of the previous three). Coefficients for the interaction between treatment and after the intervention are not significant.

In spite of the previous results, we do observe a significant *group effect.* For children that stayed at the nursery after the intervention, the prevalence of diarrhea and other illnesses before the intervention was lower in the treatment than in the control group. Maternal reported cases of diarrhea (in 7 days prior to the survey) among children in treatment nurseries were 5.8 percentage-point lower than in control nurseries (from a base of 14.2%). For respiratory diseases, the group effect was statistically not significant. In the case of illnesses different from diarrhea and respiratory diseases, the group effect was -4.8 percentage points. Combining all the illnesses in the 7 days prior to the survey (see table 40, column "any illness"), this effect was a reduction by 7.3 percentage-point on treatment nurseries. In section 3.3.3 we showed that drop-out children had no significant differences in characteristics and pre-treatment outcomes between treatment and control nurseries. Then, our estimate for group effect will combine pre-existing differences with any other difference from attrition, even when the last one was not significant.

Estimations in table 40 consider all children that stayed enrolled in the nursery after the intervention and reported illnesses at the baseline and follow up or only in one of them. [40]

[40] Similar results are obtained when we restrict the sample to those with no missing data in the outcome variable at baseline and follow up. Sample size drops to around 1,500 and standard errors increase slightly. The drop in the sample size is mainly related to cases of no attendance to the nursery on the interview days. In the 95% of the cases the nursery caregiver reported that the child was attending during the last 5 days.

TABLE 40

Linear Probability Model. Estimates on Child Health

	Diarrhea	Respiratory Disease	Other Illness	Any illness
treatment x after	0.0419	0.0306	0.0136	0.0229
	(0.034)	(0.057)	(0.033)	(0.054)
treatment	**-0.0577****	-0.0523	**-0.0480***	**-0.0732***
	(0.025)	(0.043)	(0.027)	(0.039)
after	-0.0373	**-0.1257*****	-0.0397	**-0.1203*****
	(0.025)	(0.040)	(0.025)	(0.038)
R-squared	0.017	0.026	0.011	0.028
N	1,600	1,606	1,605	1,602

Note: Robust standard errors reported in parenthesis with clustering at nursery level. All regressions include dummy variables for stratum and controls for age, age-squared, age-cubic and gender. *Significant at 10%, ** significant at 5% and *** significant at 1%.

TABLE 41

Effect on Depression

	Dependent Variable							
	score	score	log (score)		Box Cox - score		IHS - score	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
treatment	-1.6143	-1.6250	**-0.3805****	**-0.3864****	**-0.4626****	**-0.4686****	**-0.3483****	**-0.3531****
	(1.040)	(1.057)	(0.166)	(0.166)	(0.210)	(0.209)	(0.155)	(0.154)
Observations	140	138	140	138	140	138	140	138
Individual Controls	No	Yes	No	Yes	No	Yes	No	Yes

Note: Mean (score) = 6.0942, standard deviation (score)= 6.0976. Robust standard errors clustered at household level. All regressions include dummies for stratified areas. Individuals controls are: age and year of education (lineal and quadratic terms). In all cases we rescale the score from 1-61 instead of 0-60. * Significant at 10%, ** significant at 5%, and *** significant at 1%.

3.5.2 Effects on caregivers mental health

Table 41 presents the results of improving housing conditions on mental health from the OLS estimations. Column 1 and 2 report the treatment effect with and without controls for caregiver's age and years of education. Simple difference between caregivers in treatment and control nurseries shows a positive but not significant effect on mental health, reducing the depression score by 1.62 points (0.27 standard deviations) or 23%.

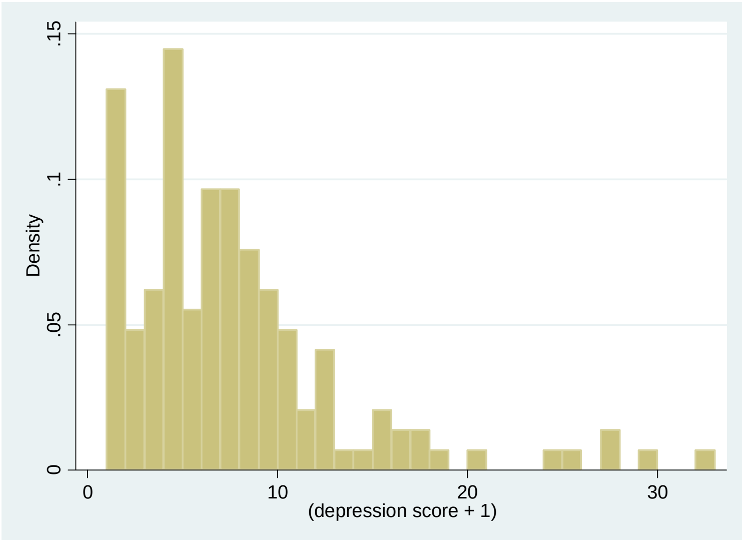
Considering that with a relatively small sample the estimates of a linear model may lack precision, that the distribution of the depression score is not normal (see 5), and the possibility of outliers, we decided to apply a non-linear transformation to circumvent this problems.

We apply three non-linear transformations: logarithmic, box cox, and inverse hyperbolic sine [41]. Before applied the transformations we rescale the score. Around 20%

[41] $IHS(z) \equiv asinh(z) = ln(z + (z^2 + 1)^{0.5})$

FIGURE 5

Distribution of Depression Score. Baseline



of the values of the depression score were zero, meaning that 20% of the caregiver interviewed after the intervention scored zero in each of the 20 items of the depression scale. To avoid loosing the information from the cases with zero score, we first rescale the score as score+1 and then apply the transformations.

Columns 3 to 6 (table 41) report the effect of the intervention on the transformed depression score. All the results are reported in terms of the rescaled score. For all the cases we find the program has a significant effect. Column 4 indicates an average treatment effect of -0.386 log point (-32%) or -0.4 standard deviations.[42] We obtain a similar results using the BC and IHS transformations. In this case, the effect is approximately a 34% and 35% reduction in the depression score, respectively. The marginal effects of the treatment on depression score, computed at the mean of the depression score for the control group after the intervention, are -2.33 and -2.43 points for the box cox and IHS transformations, respectively.[43]

In the presence of heterogeneous treatment effects it is relevant to consider the effect of the program beyond the conditional mean. In this case, a parameter of interest is the quantile treatment effect (QTE). Table 42 reports the UQTEs at the 10-th, 20-th, 30-th, 40th, 50-th, 60-th, 70-th, 80-th and 90-th percentile, following Firpo (2007), as described in section 3.4.

[42] A treatment effect of -0.386 log points corresponds to a treatment effect of (exp(-0.386)-1)*100 = 32.06 percent.

[43] For the box cox transformation, marginal effects are computed as -0.4686*6.83^(1-0.165). For the IHS is calculated as -0.3531*(-2.437^2+1)^0.5=-2.43. Note for a model $IHS(z) = X\beta + \varepsilon$, $\frac{dz}{dx} = \beta(z^2+1)^{0.5}$

TABLE 42
Unconditional Quantile Treatment Effect on Depression

	Percentile								
	10	20	30	40	50	60	70	80	90
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
treatment	0.0000	-1.0000	**-2.0000****	**-2.0000****	-1.0000	**-3.0000*****	**-3.0000***	-1.0000	1.0000
	(1.057)	(0.935)	(0.886)	(0.885)	(1.170)	(1.143)	(1.353)	(1.872)	(3.348)

Note: Unconditional quantile treatment effects based on Firpo (2007) . All regressions include dummies for stratified areas and controls for caregiver's age and year of education (lineal and quadratic terms). Sample size 138. Propensity score estimated by logit regression using an infinite bandwidth and global smoothing (lambda=1). Depression score is redefined between 1-61, in our sample the maximum value is 35. * Significant at 10%, ** significant at 5%, and *** significant at 1%.

The intervention does not tackle the determinants of depression, then we would not expect strong effects at the top of the distribution. Our results suggest that the intervention improves the mental health for caregivers in the central part of the distribution, between the 30-th and 70-th percentile. The effects are stronger above the median with a reduction by 3 points at the 60-th and 70-th percentile while they are 2 points at the 30-th and 40-th percentile. We find no effect on those who score very high or very low on the depression score (see table 42).

Conditional QTE and unconditional QTE for the non-lineal transformation on the depression score are reported in the Appendix III .

3.6 Robustness check

Attrition In section 3.3.4 we show that children that dropped out have similar characteristics irrespective of being in the control or treatment groups. In this section, we confirm our initial results assessing whether attrition is explained by treatment status and pre-treatment health outcomes and also modeling for non-random attrition.

Probability of Attrition We estimate the probability of attrition as a function of treatment nursery, pre-treatment health outcomes as well as a range of child and family characteristics. Table 43 shows that children in treatment nurseries are almost as likely to attrit than children in control nurseries and there is no significant differences in attrition based on pre-existing health outcomes. We also note that older children and children with more educated mothers are more likely to attrit.

Modeling nonrandom attrition. There are two ways in which attrition can bias the results. First, it may alter the characteristics of the sample, making it no longer representative of the original sample. This is the case if the stayers are not representative of

TABLE 43

Probability of no participation in follow up surveys

	Probability of attrition - Probit estimates				
	(1)	(2)	(3)	(4)	(5)
treatment	-0.0511	-0.0627	-0.0612	-0.0682	-0.0629
	(0.065)	(0.073)	(0.076)	(0.076)	(0.077)
age in months		-0.0831***	-0.0813***	-0.0808***	-0.0810***
		(0.024)	(0.026)	(0.026)	(0.026)
age in months ^2		0.0016***	0.0016***	0.0016***	0.0016***
		(0.000)	(0.000)	(0.000)	(0.000)
male		-0.0420	-0.0106	-0.0097	-0.0041
		(0.066)	(0.073)	(0.073)	(0.073)
mother's year of education			0.0282***	0.0276***	0.0283***
			(0.010)	(0.010)	(0.010)
mother's age			-0.0098	-0.0099	-0.0099
			(0.006)	(0.006)	(0.006)
diarrhoea, t=0					0.0338
					(0.113)
respirat. Diseases, t=0					-0.0856
					(0.068)
other illnesses, t=0					0.1037
					(0.126)
any illnesses, t=0				-0.0847	
				(0.065)	
Constant	0.1689***	0.6511	0.6475	0.7008	0.6701
	(0.060)	(0.458)	(0.526)	(0.527)	(0.530)
marginal effect for treatment	-0.0201	-0.0207	-0.0204	-0.0227	-0.0209
st. error	(0.026)	(0.024)	(0.025)	(0.025)	(0.026)
t-statistic	-0.78	-0.85	-0.81	-0.89	-0.81
N	1,958	1,860	1,561	1,561	1,560

Note: All regressions includes dummy variables for stratum. Robust standard errors reported in parenthesis with clustering at nursery level. All regressions include dummy variables for stratum. *Significant at 10%, ** significant at 5% and *** significant at 1%.

TABLE 44

Effect on Child Health from weighted least squares

	Diarrhea		Respiratory Disease		Other Illness		Any illness	
	weighted	unweighted	weighted	unweighted	weighted	unweighted	weighted	unweighted
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
treatment x after	0.0422	0.0497	0.0506	0.0754	0.0298	0.0288	0.0485	0.0710
	(0.036)	(0.034)	(0.063)	(0.061)	(0.034)	(0.036)	(0.059)	(0.058)
treatment	**-0.0595****	**-0.0614****	**-0.0928****	**-0.1029****	**-0.0611****	**-0.0585****	**-0.1099*****	**-0.1190*****
	(0.028)	(0.026)	(0.044)	(0.043)	(0.027)	(0.029)	(0.039)	(0.040)
after	-0.0462	**-0.0496***	**-0.1551*****	**-0.1631*****	-0.0429	-0.0403	**-0.1532*****	**-0.1569*****
	(0.030)	(0.027)	(0.047)	(0.045)	(0.027)	(0.027)	(0.043)	(0.041)
Observations	1,348	1,348	1,353	1,353	1,353	1,353	1,350	1,350
R-squared	0.016	0.016	0.029	0.029	0.015	0.014	0.033	0.032

Note: Odd columns report the OLS estimates using inverse probability weights from estimation of drop out on child age, gender, mother's age and years of education and dummies of health outcomes at baseline (diarrhea, respiratory disease, other illness). Even columns report OLS estimates without using weights for the sample used in the weighted estimates. Robust standard errors reported in parenthesis with clustering at nursery level. * Significant at 10%, ** significant at 5%, and *** significant at 1%.

the children interviewed at baseline. The second way that selective attrition can bias longitudinal data is by altering the covariance of variables. The underrepresentation of some groups of children leads to correlations between variables that are different than the true correlations in the original sample. In section 3.3.3 we argue that attrition does not seem be selective. Differences at baseline between drop out children in treatment and control nurseries are not statistically significant.In this section we estimate a linear panel data model under possible nonrandom attrition based on *inverse probability weighting* (IPW). Inverse probability weighting involves two steps. First, we estimate a Probit of drop out on child and family characteristics and pre-program child health outcomes. We then use the fitted probability, $\hat{p}_i$ and estimate least squares weighting by $1/\hat{p}_i$. The results we obtain from these procedure are reported in table 44.

As in section 3.5, we find a *group effect* but no impact on child health. At baseline, children in treatment nurseries experience 6 percentage point less episodes of diarrhea and other diseases, 9 percentage point less prevalence of respiratory diseases and 10 percentage point less prevalence of any illness than children in control nurseries.

As an additional check for the possibility of non-random attrition, we use Heckman's procedure to detect and correct for attrition bias. We explore whether the day of week is a good predictor for attrition. If the interview takes place in some nurseries during the weekend it is likely to observe high level of attrition. However, we we do not find any significant effect when using the day of week for the interview. Then, day of the week is not a valid instrument to estimate the drop out in the first stage of the Heckman's procedure.

TABLE 45
Linear Probability Model. Estimates on Child Health, 2009

	Diarrhea	Respiratory Disease	Other Illness	Any illness
treatment	-0.0083	-0.0129	-0.0273	-0.0351
	(0.023)	(0.040)	(0.019)	(0.041)
R-squared	0.007	0.036	0.012	0.027
N	853	858	858	853

Note: Robust standard errors reported in parenthesis with clustering at nursery level. All regressions include dummy variables for stratum and controls for age, age-squared, age-cubic, gender and pre-treatment outcomes for diarrhea, respiratory diseases and other illnesses. *Significant at 10%, ** significant at 5% and *** significant at 1%.

3.6.1 Effect on child health

Effect on child health controlling for pre-treatment illnesses In section 3.3.3 we explain that children in the control group seem to be more prompt to diseases. In this section, we control directly for preexisting differences in prevalence of illnesses between treatment and control groups instead of using the fixed effect approach to control for any time-invariant difference between groups as proposed in section 3.5.1.

Table 45 reports the effect of the intervention on child health after controlling for pre-existing differences in health outcomes. As in section 3.5.1, we find no impact on child health.

Double differences: marginal effects from probit model. Table 46 reports the marginal effects on child health from a Probit model, for the treatment group and time effects, and the interaction effect. *Group effect* is defined as the increment in the expected potential outcome Y^0 conditional on time, treatment group, and X and expressed as

$$\frac{\triangle E(Y^0|after,T,X,\lambda_s)}{\triangle T} = \Phi(\alpha+\beta+\phi after+\gamma X+\lambda_s)-\Phi(\alpha+\phi after+\gamma X+\lambda_s) \tag{13}$$

Similarly, *time effect* is defined as

$$\frac{\triangle E(Y^0|after,T,X,\lambda_s)}{\triangle after} = \Phi(\alpha+\beta T+\phi+\gamma X+\lambda_s)-\Phi(\alpha+\beta T+\gamma X+\lambda_s) \tag{14}$$

To compare the results with those reported in table 40, we compute the marginal

TABLE 46

Marginal Effects on Child Health from Probit Model

	Marginal Effects			
	Diarrhea	Respiratory Disease	Other Illness	Any Illness
treatment x after	0.025	0.029	0.005	0.025
	(0.052)	(0.057)	(0.025)	(0.055)
treatment	-0.054	-0.053	-0.050	**-0.074***
	(0.086)	(0.043)	(0.077)	(0.040)
after	-0.035	**-0.126*****	-0.043	**-0.121*****
	(0.056)	(0.040)	(0.065)	(0.043)
N	1600	1606	1605	1602

Note: Marginal effect for the coefficient of the interaction (δ) is computed as E(Y1|treat=1, after=1, x) -E(Y0|treat=1, after=1, x) that is equivalent to the cross difference of the conditional expectation of the observed outcome Y minus the cross difference of the conditional expectation of the counterfactual outcome Y (Puhani (2012). Marginal effect for treatment group defined as E(Y0|treat=1, after=0, x) - E(Y0|treat=0, after=0, x) , marginal effect for time effect = E(Y0|treat=0, after=1, x) -E(Y0|treat=0, after=0, x). Robust standar errors in parenthesis. *Significant at 10%, ** significant at 5% and *** significant at 1%.

effect for treatment *group* at the baseline ($after = 0$) based on equation 13 and the marginal effect for *time* on the control group ($T = 0$) using equation 14 . These effects are very similar to those reported in table 40.

In the probit ‘‘difference-in-differences’’ model, the treatment effect is the difference between two cross differences: it is the cross difference of the conditional expectation of the observed outcome Y minus the cross difference of the conditional expectation of the counterfactual outcome Y^0. As Puhani (2012) shows, the treatment effect δ is equal to $\frac{\Delta^2 E[Y|after,T,X,\lambda_s]}{\Delta T \Delta G} - \frac{\Delta^2 E[Y^0|after,T,X,\lambda_s]}{\Delta T \Delta G}$ and can be computed as $\Phi(\alpha+\beta+\phi+\delta+\gamma X+\lambda_s) - \Phi(\alpha+\beta+\phi+\gamma X+\lambda_s)$. The treatment effect is also not significant and smaller than the estimates from section 3.5.

Increasing the sample: Including new children In this section we investigate the effect of including children that were not attending at baseline but were attending the following year. Table 47 reports the characteristics of new children based on the follow up survey. We observe that new children in control nurseries come from poorer and less educated households than children in treatment nurseries.

We replicate table 46 using the sample of all children at baseline and the new children at follow up. As in section 3.5.1, we find improving housing condition has no impact on child health.

TABLE 47

Difference of means for new children interviewed in 2009

	Mean		N	Dif	St. Errors	t-stat
	T	C				
Family and household characteristics						
Tutor's Age	29.42	28.43	1,032	0.91	(0.586)	1.548
Tutor's Education	8.37	7.76	1,015	0.61	(0.277)	2.188
Head's Age	34.17	34.42	1,025	-0.26	(0.796)	-0.322
Head's Education	8.05	7.47	1,010	0.59	(0.285)	2.059
Poor (SISBEN 1=1)	0.95	0.98	915	-0.03	(0.016)	-1.822
Time to center	45.51	45.58	928	-0.54	(2.299)	-0.234
Time to pharmacy	12.29	12.78	929	-0.39	(0.690)	-0.563
Time to school	12.18	12.30	1,032	0.01	(0.896)	0.011
Time to hospital	34.24	33.45	928	0.94	(2.032)	0.463
Child Characteristics						
Male	0.50	0.49	1,073	0.02	(0.030)	0.793
Age in months	38.88	38.54	1,064	0.56	(0.797)	0.699
Health Service						
health insurance =1	0.97	0.96	1,014	0.01	(0.015)	0.971

TABLE 48

Linear Probability Model. Estimates on Child Health Including New Children

	Diarrhea	Respiratory Disease	Other Illness	Any illness
treatment x after	0.0216	0.0065	0.0313	0.0344
	(0.030)	(0.050)	(0.028)	(0.049)
treatment	-0.0256	-0.0361	-0.0432**	-0.0551*
	(0.018)	(0.033)	(0.021)	(0.032)
after	-0.0045	-0.0951***	-0.0342*	-0.0953***
	(0.023)	(0.034)	(0.019)	(0.035)
R-squared	0.010	0.020	0.006	0.020
N	3,652	3,662	3,660	3,654

Note: Robust standard errors reported in parenthesis with clustering at nursery level. All regressions include dummy variables for stratum and controls for age, age-squared, age-cubic and gender. *Significant at 10%, ** significant at 5% and *** significant at 1%.

3.7 Conclusions and final remarks

In this paper we investigate the effect of improving toilets, and floor and wall quality on adult mental health using a randomized experiment in the city of Cartagena, Colombia. We explore the impact of the intervention on health outcomes for children attending community nurseries (the target of the intervention). When considering diarrhea, respiratory diseases, and other illnesses, we do not find in the available data evidence of a statistically significant effect on these health outcomes. However, these results need to be interpreted with caution since, as previously discussed, we observe a high level of attrition in these subjects (61% of the initial sample). Attrition appears to be not selective relative to the intervention, affecting both control and treatment groups in the same way, and thus preserving the random nature of the group composition for the purpose of the experiment. Yet, despite this non-selective nature of the attrition, the reduction of the sample size reduces the statistical power to detect significant effects. We consider that this reduction in sample size could be particularly detrimental here since we use indirect measures (diarrhea) for our outcome of interest (parasitic infections) and, as a result, the measured coefficient would tend to be of lower statistical significance than in the case where we dispose of a direct measure of our outcome of interest.We consider that these two aspects could have played a significant role in driving these results, so they should be consider with caution.

We find evidence that improving housing conditions has a positive effect on adult's mental health of 0.3 to 0.4 standard deviations, as measured by depression scores (adjusted to account for the non-normality of the distribution and the presence of outliers). In particular, we find the effect to be significant from the 30th to the 70th percentile, indicating that the intervention had a positive effect on mental health for subjects in the middle of the distribution. We did not find a statistically significant effect for subjects initially ranking very high or very low in the measured depression score. We performed such decomposition of the results estimating the unconditional quantile treatment effects on the depression score in the spirit of Firpo (2007).

Appendix III

Table 49 shows the estimates for conditional quantile treatment effect presented in section 3.4. Estimates for CQTEs are based on the quantile regression estimators proposed by Koenker and Bassett (1978) but considering analytical standard errors that are consistent also in the case of heteroskedasticity as detailed in Frölinch and Melly (2010).

TABLE 49
Conditional Quantile Treatment Effects on Depression

	Percentile								
	10	20	30	40	50	60	70	80	90
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
A. SCORE									
treatment	0.0000	-0.2231	-0.9316	-1.8837	**-2.3416****	**-2.2102***	**-2.0518***	-1.4317	-1.4953
	(1.621)	(1.211)	(1.172)	(1.180)	(1.177)	(1.171)	(1.094)	(1.242)	(1.574)
B. LOG SCORE									
treatment	0.0000	-0.1967	**-0.6278****	**-0.7974*****	**-0.5921****	-0.3898	-0.3143	-0.1914	-0.2146
	(0.445)	(0.318)	(0.285)	(0.292)	(0.292)	(0.289)	(0.267)	(0.237)	(0.226)
C. BC SCORE									
treatment	0.0000	-0.1969	**-0.6725****	**-0.9052****	**-0.7558****	-0.5377	-0.4102	-0.2756	-0.3013
	(0.518)	(0.362)	(0.348)	(0.361)	(0.356)	(0.356)	(0.332)	(0.316)	(0.272)
D. IHS SCORE									
treatment	0.0000	-0.1969	**-0.6725***	**-0.9052****	**-0.7558****	-0.5377	-0.4102	-0.2756	-0.3013
	(0.518)	(0.362)	(0.348)	(0.361)	(0.356)	(0.356)	(0.332)	(0.316)	(0.272)

Note: Conditional quantile treatment effects based on Koenker and Bassett (1978). Analitical standar errors (Frölinch and Melly, 2010). All regressions include dummies for stratified areas and controls for caregiver's age and year of education (lineal and quadratic terms). Sample size 138. Depression score is redefined between 1-61, in our sample the maximum value is 35. * Significant at 10%, ** significant at 5%, and *** significant at 1%.

Table 50 reports the marginal effects of housing upgrades on mental health using the transformed variables. In all cases, we find a reduction in the score in the central part of the distribution. The reduction is stronger on the lower quantiles.

TABLE 50

Unconditional Quantile Treatment Effect

	Percentile								
	10	20	30	40	50	60	70	80	90
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
A. LOG SCORE									
treatment	0.0000	-0.6931	**-1.0986*****	**-0.6931***	-0.2231	**-0.5596****	**-0.4055***	-0.1054	0.0741
	(0.326)	(0.485)	(0.282)	(0.404)	(0.275)	(0.256)	(0.225)	(0.238)	(0.261)
B. BC SCORE									
treatment	0.0000	**-0.7344***	**-1.2046*****	**-0.8234***	-0.2858	**-0.7371****	**-0.5637***	-0.1528	0.1139
	(0.384)	(0.415)	(0.347)	(0.376)	(0.347)	(0.329)	(0.298)	(0.329)	(0.392)
C. IHS SCORE									
treatment	0.0000	-0.5623	**-0.9371*****	**-0.6511****	-0.2177	**-0.5494****	**-0.4017***	-0.1048	0.0739
	(0.287)	(0.346)	(0.240)	(0.303)	(0.264)	(0.246)	(0.218)	(0.232)	(0.262)

Note: Unconditional quantile treatment effects based on Firpo (2007) . All regressions include dummies for stratified areas and controls for caregiver's age and year of education (lineal and quadratic terms). Sample size 138. Propensity score estimated by logit regression using an infinite bandwidth and global smoothing (lambda=1). Depression score is redefined between 1-61, in our sample the maximum value is 35. * Significant at 10%, ** significant at 5%, and *** significant at 1%.

References

[1] Akerlof, George A., and Rachel E. Kranton. 2002. "Identity and Schooling: Some Lessons for the Economics of Education." Journal of Economic Literature, 40(4): 1167-1201.

[2] Alderman, Harold, John Hoddinott, and Bill Kinsey. 2006. "Long term consequences of early childhood malnutrition." Oxford Economic Papers 58 (3) (July 1): 450 -474.

[3] Alexander, Greg R., and Carol C. Korenbrot. 1995. "The Role of Prenatal Care in Preventing Low Birth Weight." The Future of Children 5 (1) (April 1): 103-120.

[4] Almond, Douglas, and Janet Currie. 2011. "Chapter 15 - Human Capital Development before Age Five." In Handbooks in Economics Handbook of Labor Economics, Vol 4B, ed. Orley Ashenfelter and David Card, 4, Part B:1315 - 1486. Handbook of Labor Economics. Elsevier.

[5] Angelucci, Manuela, and Giacomo De Giorgi. 2009. 'Indirect Effects of an Aid Program: How Do Cash Transfers Affect Ineligibles' Consumption?' The American Economic Review 99 (1): 486-508.

[6] Angrist, J.D., Lang, K., 2004. Does school integration generate peer effects? Evidence from Boston's Metco program. Am. Econ. Rev. 94, 1613-1634.

[7] Arbuckle, Tye E., and Lowell E. Sever. 2008. "Pesticide Exposures and Fetal Death: A Review of the Epidemiologic Literature". Research-article.

[8] Argys, L.M., Rees, D.I., 2008. Searching for peer group effects: A test of the contagion hypothesis. Rev. Econ. Stat. 90 (3), 442-458.

[9] Artadi, Elsa 2005. Going into labor: Earnings vs infant survival in rural Africa, MIMEO.

[10] Attanasio, O. *et al*. 2004. Baseline Report on the Evaluation of Familias en Acción. www.ifs.org.uk/publications.php?publication_id=2087.

[11] Attanasio, Orazio, and Marcos Vera-Hernández. 2006. "Medium and Long Run Effects of Nutrition and Child Care: Evaluation of a Community Nursery program in Rural Colombia". Institute for Fiscal Studies.

[12] Attanasio, Orazio, Emla Fitzsimons, Ana Gomez, Martha Isabel Gutiérrez, Costas Meghir, and Alice Mesnard. 2010. Children's Schooling and Work in the Presence of a Conditional Cash Transfer Program in Rural Colombia. Economic Development and Cultural Change 58 (2) (January 1): 181-210.

[13] Attanasio, Orazio, Erich Battistin, Emla Fitzsimons, Alice Mesnard and Marcos Vera-Hernandez. 2005. How effective are conditional cash transfers? Evidence from Colombia, IFS Briefing Notes , BN54.

[14] Baird, Sarah, Jed Friedman, and Norbert Schady. 2011. "Aggregate Income Shocks and Infant Mortality in the Developing World." Review of Economics and Statistics 93 (3) (September 19): 847-856.

[15] Banerjee, Abhijit, Esther Duflo, Gilles Postel-Vinay, and Tim Watts. 2010. 'Long-Run Health Impacts of Income Shocks: Wine and Phylloxera in Nineteenth-Century France'. Review of Economics and Statistics 92 (4) (June 4): 714–728.

[16] Bardají, Azucena, Betuel Sigauque, Sergi Sanz, María Maixenchs, Jaume Ordi, John J Aponte, Samuel Mabunda, Pedro L Alonso, and Clara Menéndez. 2011. "Impact of Malaria at the End of Pregnancy on Infant Mortality and Morbidity." Journal of Infectious Diseases 203 (5) (March 1): 691 -699.

[17] Barreca, Alan I. 2010. "The Long-Term Economic Impact of In Utero and Postnatal Exposure to Malaria." Journal of Human Resources 45 (4): 865 -892.

[18] Becker, Gary S., and Nigel Tomes. 1986. "Human Capital and the Rise and Fall of Families." Journal of Labor Economics 4 (3) (July 1): S1-S39.

[19] Behrman, Jere R., and Mark R. Rosenzweig. 2004. "Returns to Birthweight." Review of Economics and Statistics 86 (2) (May 1): 586-601.

[20] Berkowitz, Gertrud S., James G. Wetmur, Elena Birman-Deych, Josephine Obel, Robert H. Lapinski, James H. Godbold, Ian R. Holzman, and Mary S. Wolff. 2003. "In Utero Pesticide Exposure, Maternal Paraoxonase Activity, and Head Circumference." Environmental Health Perspectives 112 (November 18): 388-391.

[21] Bhalotra, Sonia, and Arthur van Soest. 2008. "Birth-spacing, fertility and neonatal mortality in India: Dynamics, frailty, and fecundity." Journal of Econometrics 143 (2) (April): 274-290.

[22] Bhalotra, Sonia. 2010. "Fatal fluctuations? Cyclicality in infant mortality in India." Journal of Development Economics 93 (1) (September): 7-19.

[23] Blau, David M. 1999. "The Effect of Income on Child Development." Review of Economics and Statistics 81 (2): 261-276.

[24] Bleakley, Hoyt. 2010. "Malaria Eradication in the Americas: A Retrospective Analysis of Childhood Exposure." American Economic Journal: Applied Economics 2 (2) (April): 1-45.

[25] Blume, Lawrence E., William A. Brock, Steven N. Durlauf, and Yannis M. Ioannides. 2011. 'Chapter 18 - Identification of Social Interactions'. In Handbook of Social Economics, edited by Alberto Bisin and Matthew O. Jackson Jess Benhabib, Volume 1:853–964. North-Holland.

[26] Bobonis, G.J., Finan, F., 2009. Neighborhood peer effects in secondary school enrollment decisions. Rev. Econ. Stat. 91 (4), 695-716.

[27] Bobonis, Gustavo J, and Frederico Finan. 2009. Neighborhood Peer Effects in Secondary School Enrollment Decisions. Review of Economics and Statistics 91 (4) (November 1): 695-716.

[28] Bond, Lyndal, Ade Kearns, Phil Mason, Carol Tannahill, Matt Egan, and Elise Whitely. 2012. 'Exploring the Relationships Between Housing, Neighbourhoods and Mental Wellbeing for Residents of Deprived Areas'. BMC Public Health 12 (January 18): 48. doi:10.1186/1471-2458-12-48.

[29] Bonzini, Matteo, David Coggon, and Keith T Palmer. 2007. "Risk of prematurity, low birthweight and pre?eclampsia in relation to working hours and physical activities: a systematic review." Occupational and Environmental Medicine 64 (4) (April): 228-243.

[30] Bonzini, Matteo, David Coggon, Keith Godfrey, Hazel Inskip, Sarah Crozier, and Keith Palmer. 2009. "Occupational physical activities, working hours and outcome of pregnancy: findings from the Southampton women's survey." Occupational and environmental medicine 66 (10) (October): 685-690.

[31] Bramoullé, Y., Djebbari, H., Fortin, B., 2009. Identification of peer effects through social networks. J. Econom. 150, 41-55.

[32] Brisbois, Ben. 2011. "Pesticides, People and Power in Ecuador's Banana Industry: Participatory Epidemiology and Political Ecology Approaches to Occupational Health and Safety." Antipode 43 (1) (January 1): 176-179.

[33] Brock, W., and S. Durlauf. 2001. 'Chapter 54 Interactions-based Models'. In , Volume 5:3297-3380. Elsevier.

[34] Brock, W., and S. Durlauf. 2007. 'Identification of Binary Choice Models with Social Interactions'. Journal of Econometrics 140 (1): 52 - 75

[35] Brock, W.A., Durlauf, S.N., 2001. Interactions-based models. In: Heckman, J., Leamer, E. (Eds.), Handbook of Econometrics. vol. 5. Elsevier Science, pp. 3297-3380.

[36] Bunyavanich, Supinda, Christopher P. Landrigan, Anthony J. McMichael, and Paul R. Epstein. 2003. "The Impact of Climate Change on Child Health." Ambulatory Pediatrics 3 (1): 44-52.

[37] Burdorf, A, T Brand, V W Jaddoe, A Hofman, J P Mackenbach, and E A P Steegers. 2011. "The effects of work-related maternal risk factors on time to pregnancy, preterm birth and birth weight: the Generation R Study." Occupational and Environmental Medicine 68 (3) (March 1): 197 -204.

[38] Burgess, Robin, Olivier Deschenes, Dave Donaldson and Michael Greenstone. 2011. Weather and Death. Unpublished manuscript.

[39] Camacho, A. and E. Conover, 2009. Manipulation of Social Program Eligibility: Detection, Explanations and Consequences for Empirical Research. CEDE (19). ISSN 1657-7191.

[40] Card, David, and Alan B. Krueger. 1992. Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States. The Journal of Political Economy 100 (1) (February): 1-40.

[41] Case, A. and Katz, F. 1991.The Company You Keep: The Effect of Family and Neighborhood on Disadvantaged Youths, NBER working paper no. 3705.

[42] Cattaneo, Matias D, Sebastian Galiani, Paul J Gertler, Sebastian Martinez, and Rocio Titiunik. 2009. 'Housing, Health, and Happiness'. American Economic Journal: Economic Policy 1 (1) (January): 75-105.

[43] Chay, Kenneth Y., and Michael Greenstone. 2003. "The Impact of Air Pollution on Infant Mortality: Evidence from Geographic Variation in Pollution Shocks Induced by a Recession*." Quarterly Journal of Economics 118 (3): 1121-1167.

[44] Cipollone, Piero, and Alfonso Rosolia, "Social Interactions in High School: Lessons from an Earthquake," American Economic Review 97 (2007) 948-965.

[45] Coady, D. and R.L. Harris. 2001. Evaluating Transfer Programs within a General Equilibrium Framework. FCND Discussion Paper 110. IFPRI, Washington, DC.

[46] Currie, Janet, and Johannes F. Schmieder. 2008. "Fetal Exposure to Toxic Releases and Infant Health." National Bureau of Economic Research Working Paper Series No. 14352 (September).

[47] Currie, Janet. 2000. "Chapter 19 Child health in developed countries." In , Volume 1, Part 2:1053-1090. Elsevier.

[48] Currie, Janet. 2011. "Inequality at Birth: Some Causes and Consequences." National Bureau of Economic Research Working Paper Series No. 16798 (February).

[49] Cutler, David M, Felicia Knaul, Rafael Lozano, Oscar Mendez, and Beatriz Zurita. 2002. "Financial crisis, health outcomes and ageing: Mexico in the 1980s and 1990s." Journal of Public Economics 84 (2) (May): 279-303.

[50] Cutler, David M., Grant Miller, and Douglas M. Norton. 2007. "Evidence on Early-Life Income and Late-Life Health from America's Dust Bowl Era." Proceedings of the National Academy of Sciences of the United States of America 104 (33): 13244-13249.

[51] Cutler, David, Angus Deaton, and Adriana Lleras-Muney. 2006. "The Determinants of Mortality." The Journal of Economic Perspectives 20 (3) (July 1): 97-120.

[52] Cutler, David, Winnie Fung, Michael Kremer, Monica Singhal, and Tom Vogl. 2007. "Mosquitoes: The Long-term Effects of Malaria Eradication in India." National Bureau of Economic Research Working Paper Series No. 13539.

[53] Cutler, David, Winnie Fung, Michael Kremer, Monica Singhal, and Tom Vogl. 2010. "Early-life Malaria Exposure and Adult Outcomes: Evidence from

Malaria Eradication in India." American Economic Journal: Applied Economics 2 (2) (April): 72-94.

[54] Dehejia, Rajeev, and Adriana Lleras-Muney. 2004. "Booms, Busts, and Babies' Health*." Quarterly Journal of Economics 119 (3): 1091-1130.

[55] Desai, Meghna, Feiko O ter Kuile, François Nosten, Rose McGready, Kwame Asamoa, Bernard Brabin, and Robert D Newman. 2007. "Epidemiology and Burden of Malaria in Pregnancy". The Lancet Infectious Diseases 7 (2) (February): 93-104.

[56] Deschênes, Olivier, and Michael Greenstone. 2007a. "The Economic Impacts of Climate Change: Evidence from Agricultural Output and Random Fluctuations in Weather." The American Economic Review 97 (1) (March 1): 354-385.

[57] Deschênes, Olivier, and Michael Greenstone. 2007b. "Climate Change, Mortality, and Adaptation: Evidence from Annual Fluctuations in Weather in the US." National Bureau of Economic Research Working Paper Series No. 13178 (June).

[58] Deschênes, Olivier, Michael Greenstone, and Jonathan Guryan. 2009. "Climate Change and Birth Weight." American Economic Review 99 (2) (May): 211-217.

[59] Devoto, Florencia, Esther Duflo, Pascaline Dupas, William Parienté, and Vincent Pons. 2012. "Happiness on Tap: Piped Water Adoption in Urban Morocco". American Economic Journal: Economic Policy 4(4): 68-99.

[60] Domingues, Marlos Rodrigues, Alicia Matijasevich, and Aluísio J.D. Barros. 2009. "Physical Activity and Preterm Birth." Sports Medicine 39 (November): 961-975.

[61] Duflo, E. 2000. Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment. National Bureau of Economic Research Working Paper Series No. 7860 (August).

[62] Duflo, E., and Saez, E. 2003. The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment, Quarterly Journal of Economics 118, 815-842.

[63] Duflo, E., Dupas, P., & Kremer, M. 2011. Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya. American Economic Review , 101 (5), 1739-1774.

[64] Duflo, Esther. 2001. Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment. The American Economic Review 91 (4) (September): 795-813.

[65] Durlauf, Steven N. 2004. 'Chapter 50 Neighborhood Effects'. In Cities and Geography, edited by J. Vernon Henderson and Jacques-Francois Thisse, 4:2173 - 2242. Handbook of Regional and Urban Economics. Elsevier.

[66] Edmonds, Eric V., and Norbert Schady. 2009. Poverty Alleviation and Child Labor. National Bureau of Economic Research Working Paper Series No. 15345 (September).

[67] Edward, M. and M. Kremer, "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities," Econometrica 72:1 (2004), 159-217.

[68] Epple, D., & Romano, R. E. 2011. Peer Effects in Education: A Survey of the Theory and Evidence. Handbook of Social Economics, 1B, 1053-1163.

[69] Escriba-Agüir, Vicenta, Santiago Perez-Hoyos, and Marie-Josephe Saurel-Cubizolles. 2001. "Physical load and psychological demand at work during pregnancy and preterm birth." International Archives of Occupational and Environmental Health 74 (8): 583-588.

[70] Esrey, S. A., J. B. Potash, L. Roberts, and C. Shiff. 1991. "Effects of Improved Water Supply and Sanitation on Ascariasis, Diarrhoea, Dracunculiasis, Hookworm Infection, Schistosomiasis and Trachoma." Bulletin of the World Health Organization, 69(5): 609-21.

[71] Fafchamps, M. and Lund, S., (2003), "Risk-sharing Networks in Rural Philippines", Journal of Development Economics, 71(2): 261-87.

[72] Ferreira, F. H. G., and N. Schady. 2009. "Aggregate Economic Shocks, Child Schooling, and Child Health." The World Bank Research Observer 24 (July 8): 147-181.

[73] Firpo, Sergio. 2007. 'Efficient Semiparametric Estimation of Quantile Treatment Effects'. Econometrica 75 (1) (January 1): 259-276. doi:10.1111/j.1468-0262.2007.00738.x.

[74] Firpo, Sergio. 2007. 'Efficient Semiparametric Estimation of Quantile Treatment Effects'. Econometrica 75 (1) (January 1): 259-276.

[75] Fishback, Price V., Michael R. Haines, and Shawn Kantor. "Births, Deaths, and New Deal Relief during the Great Depression." The Review of Economics and Statistics 89,1 (February 2007): 1-14.

[76] Foster, Andrew D. and Mark R. Rosenzweig (1995), "Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture", Journal of Political Economy 103(6): 1176-1209.

[77] Frolich, Markus, and Blaise Melly. 2010. 'Estimation of Quantile Treatment Effects with Stata'. Stata Journal 10 (3): 423-457.

[78] Galiani, Sebastian, Paul J. Gertler, and Ernesto Schargrodsky. 2005. "Water for Life: The Impact of the Privatization of Water Services on Child Mortality." Journal of Political Economy, 113(1): 83-120.

[79] Glaeser, E., Scheinkman, J., 2001. Measuring social interactions. In: Durlauf, S.N., Young, H.P. (Eds.), Social dynamics, Washington, D.C. : Brookings Institution Press, Cambridge, Mass : MIT Press, 2001. pp. 83-132.

[80] Glaeser, E.L., Sacerdote, B., Scheinkman, J., 2003. The social multiplier. J. Eur. Econ. Assoc. 1, 345-353.

[81] Goldsmith-Pinkham, P. and G. Imbens. 2013. Social Networks and the Identification of Peer Effects. Journal of Business and Economic Statistics.

[82] González-Forteza, Catalina, José Alberto Jiménez-Tapia, Luciana Ramos-Lira, and Fernando A. Wagner. 2008. 'Aplicación De La Escala De Depresión Del Center of Epidemiological Studies En Adolescentes De La Ciudad De México'. Salud Pública De México 50 (4) (August): 292-299.

[83] Greenstone, Michael, and Rema Hanna. 2011. "Environmental Regulations, Air and Water Pollution, and Infant Mortality in India." National Bureau of Economic Research Working Paper Series No. 17210 (July).

[84] Haines, Michael R. 2010. "Inequality and Infant and Childhood Mortality in the United States in the Twentieth Century." National Bureau of Economic Research Working Paper Series No. 16133 (June).

[85] Handal, Alexis, and Sioban Harlow. 2009. "Employment in the Ecuadorian cut-flower industry and the risk of spontaneous abortion." BMC International Health and Human Rights 9 (1): 25.

[86] Hanushek, E.A., Kain, J.F., Markman, J.M., Rivkin, S.G., 2003. Does peer ability affect student achievement? J. Appl. Econom. 18, 527-544.

[87] Hanushek, E.A., Kain, J.F., Rivkin, S.G., 2009. New evidence about Brown v. Board of Education: The complex effects of school racial composition on achievement. J. Labor Econ. 27 (3), 349-383.

[88] Hatch, Maureen, Bu-Tian Ji, Xiao Ou Shu, and Mervyn Susser. 1997. "Do Standing, Lifting, Climbing, or Long Hours of Work during Pregnancy Have an Effect on Fetal Growth?" Epidemiology 8 (5): 530-536.

[89] Hatfield, J. L., K. J. Boote, B. A. Kimball, L. H. Ziska, R. C. Izaurralde, D. Ort, A. M. Thomson, and D. Wolfe. 2011. "Climate Impacts on Agriculture: Implications for Crop Production." Agron. J. 103 (2) (March): 351-370.

[90] Hay, S.; Guerra, C.; Gething, P.; Patil, A.; Tatem, A.; Noor, A.; Kabaria, C.; Manh, B.; Elyazar, I.; Brooker, S.; Smith, D.; Moyeed, R. & Snow, R. 2009. "A world malaria map: Plasmodium falciparum endemicity in 2007" PLoS Medicine, 6 (3).

[91] Heckman, James J., and V. Joseph Hotz. 1989. Choosing Among Alternative Nonexperimental Methods for Estimating the Impact of Social Programs: The Case of Manpower Training. Journal of the American Statistical Association 84 (408) (December): 862-874.

[92] Imbens, Guido W. 2004. 'Nonparametric Estimation of Average Treatment Effects Under Exogeneity: A Review'. The Review of Economics and Statistics 86 (1) (February 1): 4-29.

[93] Imberman, S., Kugler, A., Sacerdote, B., 2009. Katrina's children: Evidence on the structure of peer effects from hurricane evacuees. NBER Working Paper No. 15291.

[94] Jayachandran, Seema, and Adriana Lleras-Muney. 2009. "Life Expectancy and Human Capital Investments: Evidence from Maternal Mortality Declines*." Quarterly Journal of Economics 124 (1) (February 1): 349-397.

[95] Katz, Lawrence F., Jeffrey R. Kling, and Jeffrey B. Liebman. 2001. Moving to Opportunity in Boston: Early Results of a Randomized Mobility Experiment*. Quarterly Journal of Economics 116 (2) (May 1): 607-654.

[96] Kjell Doksum. Empirical probability plots and statistical inference for nonlinear models in the two-sample case. Ann. Statist., 2(2):267-277, 1974. URL http://projecteuclid.org/euclid.aos/1176342662.

[97] Kremer, M., Miguel, E., Thornton, R., 2009. Incentives to learn. Rev. Econ. Stat. 91 (3), 437-456.

[98] Kudamatsu, Masayuki, Persson, Torsten and Strömberg, David, Weather and Infant Mortality in Africa (November 2012). CEPR Discussion Paper No. DP9222. Available at SSRN: http://ssrn.com/abstract=2210191

[99] Lalive, Rafael, and M. Alejandra Cattaneo. 2009. Social Interactions and Schooling Decisions. Review of Economics and Statistics 91 (3): 457-477.

[100] Lima, M, S Ismail, A Ashworth, and S S Morris. 1999. "Influence of heavy agricultural work during pregnancy on birthweight in northeast Brazil." International Journal of Epidemiology 28 (3) (June): 469-474.

[101] Luo, Qunying. 2011. "Temperature thresholds and crop production: a review." Climatic Change (February).

[102] Maccini, Sharon, and Dean Yang. 2009. "Under the Weather: Health, Schooling, and Economic Consequences of Early-Life Rainfall." American Economic Review 99 (3) (June): 1006-1026.

[103] Maluccio, J. and R. Flores. 2005. Impact Evaluation of the Pilot Phase of the Nicaraguan Red de Protección Social. Research Report 141, IFPRI, Washington, D.C.

[104] Manski, C. F. 1993. Identification of Endogenous Social Effects: The Reflection Problem. Review of Economic Studies, 60(3), 531-542.

[105] Matamoros, David and Peter A Vanrolleghem. 2001. "Pesticide Assessment Of The Banana Sector In An Ecuadorian Watershed." MIMEO. Escuela Superior Politécnica del Litoral,

[106] Mayer, S. 1997. What Money Can't Buy: Family Income and Children's Life Chances. Cambridge, MA: Harvard Univ. Press.

[107] Miller, Grant, and B. Piedad Urdinola. 2010. "Cyclicality, Mortality, and the Value of Time: The Case of Coffee Price Fluctuations and Child Survival in Colombia." Journal of Political Economy 118 (1) (February 1): 113-155.

[108] Minnesota Population Center. Integrated Public Use Microdata Series, International: Version 6.1 [Machine-readable database]. Minneapolis: University of Minnesota, 2011.

[109] Moffit, Robert, John Fitzgerald, and Peter Gottschalk. 1999. 'Sample Attrition in Panel Data: The Role of Selection on Observables'. Annals of Economics and Statistics / Annales d'Économie Et De Statistique (55/56) (September 1): 129-152.

[110] Moffit. 2001. 'Policy Interventions, Low-level Equilibria, and Social Interactions' in Steven N. Durlauf and H. Peyton Young (eds.), Social Dynamics (Washington, DC: Brookings Institution Press) pp. 45-82.

[111] Munshi, K., and Myaux, J. 2006. Social Norms and the Fertility Transition, Journal of Development Economics 80 (2006), 1-38.

[112] National Administrative Department of Statistics. Census 2005.

[113] Neidell, Matthew, and Jane Waldfogel. 2008. Cognitive and Non-Cognitive Peer Effects in Early Education. National Bureau of Economic Research Working Paper Series No. 14277 (August).

[114] Norbert Schady, and Maria Caridad Araujo. 2008. Cash Transfers, Conditions, and School Enrollment in Ecuador. Economía 8 (2): 43-70.

[115] Paxson, Christina H. 1992. "Using Weather Variability to Estimate the Response of Savings to Transitory Income in Thailand." The American Economic Review 82 (1) (March): 15-33.

[116] Paxson, Christina, and Norbert Schady. 2005. "Child Health and Economic Crisis in Peru." The World Bank Economic Review 19 (2) (January 1): 203 -223. doi:10.1093/wber/lhi011.

[117] Pitt, Mark M., Mark R. Rosenzweig, and Donna M. Gibbons. 1993. The Determinants and Consequences of the Placement of Government Programs in Indonesia. The World Bank Economic Review 7 (3) (September): 319-348.

[118] Price, Joseph. 2008. "Parent-Child Quality Time." Journal of Human Resources 43 (1) (December 21): 240 -265.

[119] Pritchett, Lant, and Lawrence H. Summers. 1996. "Wealthier is Healthier." The Journal of Human Resources 31 (4) (October 1): 841-868.

[120] Puhani, Patrick A. 2012. 'The Treatment Effect, the Cross Difference, and the Interaction Term in Nonlinear "difference-in-differences" Models'. Economics Letters 115 (1) (April): 85-87.

[121] Rabassa, Mariano, Skoufias, Emmanuel, and Jacoby, Hanan G., 2012. "Weather and child health in rural Nigeria," Policy Research Working Paper Series 6214, The World Bank.

[122] Rayco-Solon, Pura, Anthony J Fulford, and Andrew M Prentice. 2005. "Differential effects of seasonality on preterm birth and intrauterine growth restriction in rural Africans." The American Journal of Clinical Nutrition 81 (1) (January 1): 134 -139.

[123] Roberts EH, Summerfield RJ (1987) Measurement and prediction of flowering in annual crops. In: Atherton JG (ed) Manipulation of flowering. Butterworths, London, pp 17–50.

[124] Rogerson, Stephen J., Victor Mwapasa, and Steven R. Meshnick. 2007. "Malaria in Pregnancy: Linking Immunity and Pathogenesis to Prevention". Am J Trop Med Hyg 77: 14-22.

[125] Rosenzweig, Cynthia, Ana Iglesias, X.B. Yang, Paul R. Epstein, and Eric Chivian. 2001. "Climate Change and Extreme Weather Events; Implications for Food Production, Plant Diseases, and Pests." Global Change & Human Health 2 (2): 90-104.

[126] Rosenzweig, Mark R., and Hans P. Binswanger. 1993. "Wealth, Weather Risk and the Composition and Profitability of Agricultural Investments." The Economic Journal 103 (416) (January): 56-78.

[127] Rosenzweig, Mark R., and T. Paul Schultz. 1983. "Estimating a Household Production Function: Heterogeneity, the Demand for Health Inputs, and Their Effects on Birth Weight." The Journal of Political Economy 91 (5) (October): 723-746.

[128] Rötter, R., and S.C. van de Geijn. 1999. "Climate Change Effects on Plant Growth, Crop Yield and Livestock." Climatic Change 43 (4): 651-681.

[129] Ruhm, Christopher J. 2000a. "Are Recessions Good for Your Health?*." Quarterly Journal of Economics 115 (2) (May 1): 617-650.

[130] Ruhm, Christopher J. 2000b. "Parental leave and child health." Journal of Health Economics 19 (6): 931-960.

[131] Ruhm, Christopher J. 2004. "Macroeconomic Conditions, Health and Mortality." National Bureau of Economic Research Working Paper Series No. 11007 (December).

[132] Sacerdote, B. 2001. Peer effects with random assignment: Results for Dartmouth roommates. Q. J. Econ. 116, 681-704.

[133] Sacerdote, B. 2011. Peer Effects in Education: How Might They Work, How Big Are They and How Much Do We Know Thus Far? Handbook of Economics of Education, 3, 249-277.

[134] Schultz, T. Paul. 1985. "Changing World Prices, Women's Wages, and the Fertility Transition: Sweden, 1860-1910." The Journal of Political Economy 93 (6) (December): 1126-1154.

[135] Skoufias, Emmanuel, B. Essama-Nssah, and Roy Katayama. 2011. "Too Little Too Late: Welfare Impacts of Rainfall Shocks in Rural Indonesia." Policy Research Working Paper Series No. 5615. World Bank.

[136] Steketee, RW, BL Nahlen, ME Parise, and C Menendez. 2001. 'The Burden of Malaria in Pregnancy in Malaria-endemic Areas'. Am J Trop Med Hyg 64 (1_suppl) (January 1): 28-35.

[137] Stillerman, Karen Perry, Donald R. Mattison, Linda C. Giudice, and Tracey J. Woodruff. 2008. "Environmental Exposures and Adverse Pregnancy Outcomes: A Review of the Science." Reproductive Sciences 15 (7): 631 -650.

[138] Thomson, Hilary, Sian Thomas, Eva Sellstrom, and Mark Petticrew. 2009. 'The Health Impacts of Housing Improvement: A Systematic Review of Intervention Studies From 1887 to 2007'. American Journal of Public Health 99 (S3) (November): S681-S692. doi:10.2105/AJPH.2008.143909.

[139] Townsend, Robert. 1994. "Risk and Insurance in Village India." Econometrica, 62(3): 539-91.

[140] Tiwari, Sailesh, and Hanan G. Jacoby. 2013. ‘Monsoon Babies: Rainfall Shocks and Child Nutrition in Nepal’. SSRN Scholarly Paper ID 2241953. Rochester, NY: Social Science Research Network.

[141] Wheeler TR, Craufurd PQ, Ellis RH, Porter JR, Vara Prasad PV (2000) Temperature variability and the yield of annual crops. Agric Ecosyst Environ 82:159–167

[142] WHO. 2003. Antenatal care in developing countries: Promises, achievements and missed opportunities: an analysis of trends, levels and differentials, 1990-2001.

[143] Wooldridge, J. 2002. Econometric analysis of cross section and panel data.

[144] Zimmerman, D.J., 2003. Peer effects in academic outcomes: Evidence from a natural experiment. Rev. Econ. Stat. 85, 9-23.