When fathers are gone: the consequences of paternal absence during the early years*

Julieta Vera Rueda†

Job Market Paper

November 14, 2022

Abstract

I examine the causal effect of paternal absence during the first years of life on early childhood development. I take advantage of Young Lives study, a unique panel dataset including skill outcomes of pairs of siblings with different levels of paternal involvement. With a difference in difference approach, I first compare the younger sibling exposed to the father's absence in the early years versus the older one exposed at age 5. Then, I control for the gap in skills between siblings living in households with no absence. Results show that paternal absence in the early years leads to a relative worsening in nutritional outcomes when 5 years old. The gap persists when 8 and 12 years old, with no evidence of cumulative effects.On the contrary, estimates for cognition indicate there is no negative effect of early absence. If anything, younger siblings relatively outperform their older counterparts. Finally, there is no strong evidence on differences in development when absence occurs at age 5 for the younger and 8 for the older sibling. These findings are consistent with critical windows for ability development during the first 1,000 days of life.

Keywords: Early childhood development, Fatherhood, Developing countries

JEL Codes: O15, J13

*The data used in this publication come from Young Lives, a 15-year study of the changing nature of childhood poverty in Ethiopia, India, Peru and Vietnam (www.younglives.org.uk). Young Lives is funded by UK aid from the Department for International Development (DFID). The views expressed here are those of the author(s). They are not necessarily those of Young Lives, the University of Oxford, DFID or other funders.

†Paris School of Economics. E-mail: julieta.vera@psemail.eu. I am grateful to my advisor Karen Macours for her support and extensive feedback on this project. I would also like to thank Michael Kremer, Sylvie Lambert, Luc Behaghel, Oliver Vanden Eynde and participants at the Casual Friday Development Seminar and the Early childhood education and social inequalities Workshop at Paris School of Economics, as well as participants at the 100 Years of Economic Development Conference at Cornell University and Oxford Development Economics Workshop for helpful comments. This work has been funded by a French government subsidy managed by the Agence Nationale de la Recherche under the framework of the Investissements d'avenir programme reference ANR-17-EURE-001. I acknowledge the financial support of the Norface Dynamics of Inequality Across the Lifecourse (DIAL) Joint Research Programme (research Project file number 462-16-090, entitled Human capital and inequality during adolescence and working life) and the Agence Nationale de la Recherche (project ANR-17-0004-01). All remaining errors are my own.

1 Introduction

The role of parents in human capital formation has garnered substantial attention in the field of economics. This can be explained by the following facts. The early years set the foundations for an ongoing process of skill accumulation. Indeed, the first 1,000 days of life–from conception to age 2– are critical windows for anthropometric outcomes (Victora et al. 2010). Because skills matter for adult success, early childhood development has strong predictive power of outcomes such as education and wages (Black et al. 2017, Currie & Almond 2011). The process in which skills are formed is not only influenced by genetics but also by experiences faced by the child (Shonkoff et al. 2012). Moreover, experiences during the early years are mostly influenced by the home environment; but more specifically by interactions between the child and parents (Cunha & Heckman 2007). Altogether, this stresses how important parents are for future adult success.

Parenting plays a crucial role both in terms of nutrition and nurturing. Parents make fundamental choices such as food allocation and health investments, as well as offering emotional support and mental stimulation. Moreover, parent-child interactions can exacerbate or mitigate risk factors. The specific role of the mother and the father in this process is still an open question.

In the economic literature, most theoretical models of human capital formation limit the analysis to one adult and one child, with no distinction between mothers and fathers (Cunha & Heckman 2007). Others frame childcare distribution between parents in a setting of intra-household resource allocation (Becker 1991), where the relative allocation is mainly the result of a comparative advantage in caregiving. From an empirical perspective, most of the evidence comes from early childhood interventions that have mainly focused on the mother as the main provider of inputs for children's development. This is adequate as long as maternal childcare is all that matters or if caregiving provided by either of the parents is perfectly substitutable.

Fathers may have a differential effect on the development of children for two reasons. First of all, fathers may complement maternal childcare by providing particular inputs which boost those provided by the mother. Secondly, fathers could also affect specific inputs that mothers do not. Either way, beyond the inaccuracy of the mentioned models, ignoring the role of the father in skill formation could potentially limit the impact of policies aiming at reducing the development gap observed across socioeconomic groups.

This paper answers the question whether the input of the father is important during the critical period of the first 1,000 days of life. I study the extreme case in which the father is absent, trying to understand if absence has an impact on skill formation; directly and indirectly. Moreover, I test whether the impact is the same across abilities. In this way, I seek to establish a causal relationship between paternal absence during the critical period on early childhood development outcomes in four developing countries– Ethiopia, India, Peru and Vietnam. I consider early childhood as the period that goes up to age 5, as evidence suggests that some ability gaps become salient only after some periods.[1]

I contribute to the human capital literature in the following dimensions. I isolate paternal inputs from other caregivers' inputs, adding some insight on the relevance of the father on the skills production function. Secondly, I focus on the early years, highlighting the pertinence of paternal investments during the critical periods. I explore nutritional, cognitive and socio-emotional outcomes, which enables me to look

[1] For instance, Schady (2011) finds how the distribution in vocabulary scores among poor children in rural Ecuador changes during the period between 3 and 5 years old. Similarly, Rubio-Codina et al. (2015) find a widening in the average gap in cognitive and language skills after 24 months of age in Bogota, Colombia.

at how paternal absence affects each type of ability. In addition, I measure those outcomes in the short and the long run. More importantly, the methodology I use provides internal validity on the causal role of the father in early childhood development. Lastly, the evidence in four developing countries provides external validity of the results.

Finding a causal relationship between paternal absence and skills development is not an easy task. One of the main challenges is the endogeneity that goes along with children's outcomes and family decisions. One way to overcome this problem is by using siblings as counterfactuals, in which one of them experiences the departure of the father while the other does not. Intra-household comparisons would remove the effect of observable and unobservable characteristics that are common to both siblings. However, siblings could develop differently irrespectively of the changes in family structure. To overcome this issue, siblings in households with no absent fathers could also be compared. Then, if gaps in development between siblings in both types of households are indeed different from each other, it is credible that paternal absence has an impact on children's development.

I take advantage of the unique dataset provided by the Young Lives study which includes information on disadvantaged children in Ethiopia, India, Peru and Vietnam. The study follows up these children (from now on, the older sibling or index child) from early childhood for a 15 year span, collecting detailed data on development outcomes, information on parents and household and community characteristics. Five waves are used in this study: 2002 (the baseline), 2006, 2009, 2013 and 2016. The main novelty of Young Lives for this paper is the inclusion of a survey on younger siblings in the last three waves. Exploiting this rich data, my sample comprises the index child, born in 2001-02, and his/her younger brother/sister born in 2003-05. I make use of the availability of nutritional outcomes, cognitive tests and socio-emotional measures[2] of both children. Furthermore, I use the frequency with which the child sees his/her father in each round as a proxy of paternal presence. I will consider a father as absent if he does not see his child (the index child) on a daily basis or if he does, he does not live at home.

This data provides me with two sources of variation. First, I end up with two types of household, one that experiences the absence of the father in one of the rounds under study and another that does not. Second, given that absence occurs at one point in time, I exploit the chronological moment of the absence of the father and the corresponding age of each sibling when that happens. More specifically, when the older sibling is around 5 while the younger is in her/his early years. To mimic a child's counterfactual as much as possible, I compare the outcomes of pairs of siblings when they are of similar age measured at two different points in time, i.e. two consecutive waves. In other words, when assessing the gap when both are around 5 years old, I compare the older sibling in the second round (2006) and the younger in the third one (2009). Adding household fixed effects, my approach resembles a difference-in-difference methodology. The first difference is the gap within household in the development of siblings, measured when they are at a similar stage in life. The second one is the difference in the gaps between the two types of household (those that experience absence and those that do not). I analyze the short term impact on the outcomes when both are aged 5 and the long term effects when both are 8 and 12 years old.

Exploiting earlier versus later exposure to paternal absence, I find that the father's absence seems to matter more for children's outcomes when aged 5 if it occurs during the first years of life relatively to at a later stage. However, the effect of paternal input differs across abilities. On the one hand, paternal absence seems to be detrimental to nutritional outcomes, and this relative under-performance in an-

[2]Outcomes are: weight-for-age, height-for-age, Peabody Picture Vocabulary Test and self-efficacy and self-esteem indices.

thropometric measures prevails over time while no cumulative effects are found. In contrast, it seems not to be the case for cognition, where I encounter no differential performance between siblings when pooling all countries.[3] In fact coefficients are positive for Peru and Vietnam, being statistically significant for the former country. In terms of socio-emotional outcomes, estimates do not give a clear interpretation as they vary across the countries and ages at which siblings are measured.

Estimates could potentially reflect the relative importance of the father's presence earlier versus later in life, irrespective of if it happens around the 1,000 day window. Moreover, measures are in relative terms. Therefore, they could be showing the impact on the older sibling and not on the one suffering absence just after birth. In order to rule out these alternatives, I perform a second exercise in which I explore the case where absence occurs at age 5 and 8 for the younger and older sibling respectively. Indeed, I encounter that having an absent father at 5 years old is not any different to experiencing it afterward in terms of nutrition and cognitive outcomes in the short nor the long-run.

An interesting question is to understand what are the factors that are affected during paternal departure that could be explaining my results. Exploring possible mechanisms, evidence suggests that the relatively more detrimental effect on nutrition can be explained by a negative income shock. Households suffering paternal absence are more likely to report a worsening in their economic situation with respect to those that did not experience paternal departure. On the other hand, results on cognition could be potentially explained by changes in family structure. It seems that households experiencing a father's departure are less likely to have other adult members leaving the household and mothers are more likely to be replaced by other household members as main caregiver. Those adults are the ones potentially compensating lack of input by the father on cognitive skills formation. Additionally, investment decisions on children by the caregiver at home might change as a result of the departure of the father. For this reason, I also look at time use by siblings, finding that the positive (still no statistically significant) point estimate in Vietnam could be explained by more time invested in productive activities and less in leisure.

It is worth mentioning that the data used in this paper does not provide information on the reason for departure, which creates two limitations. First of all, it prevents me from making any distinctions on whether absence is due to migration, divorce, abandonment or death. Each of these alternatives may have different effects on the well-being of the child as the length and the reason why he left may have different implications. I will approach this issue by studying temporary versus more permanent absence. To do this, I will analyze households in which fathers come back the next round and those that do not. Secondly, I do not know whether fathers had left as a result of a shock, which could potentially lead to a misinterpretation of the results. If this shock has a differential effect when it occurs in the early years, then my results could be capturing the impact of this particular episode rather than paternal absence. While unable to rule out this issue, I control for the occurrence of shocks reported by the household at the year of absence. I check whether results change when controlling for baseline differences in exposure to shocks. Results are consistent to these robustness checks.

These results provide support to the existence of critical periods for paternal input in the human capital production function, especially in terms of nutritional outcomes. In the countries I study, fathers seem to matter as financial supporters. Therefore, the negative income shock related to paternal absence can have a relatively higher (long lasting) detrimental effect on nutritional development when it occurs during the early years. The dissimilar effect on skill outcomes is consistent with the idea that abilities are affected by different risk factors in a heterogeneous manner. Hence,

[3] India is excluded from this analysis as there is no information on cognitive outcomes for the younger sibling.

if paternal absence has a differential impact on risk factors, then I expect to find contrasting responses on abilities. Some factors, such as lower stimulation by the father or a worsening of the home environment, can be potentially mitigated by the mother or other family members. Others, such as the negative income shock, are less easily mitigated by relatives. The former factors–stimulation and home environment– are probably more relevant for cognition and socio-emotional outcomes, explaining the null gap (and positive for Peru and Vietnam) in cognitive development. The later could be more important for anthropometric measures, backing the results found.

The remainder of the paper is organized as follows. Section 2 goes through the most relevant literature on paternal absence and on the research done using siblings from Young Lives dataset. Section 3 provides details on the data used, how siblings were sampled and the outcomes to be measured. Section 4 goes into detail on the identification strategy and econometric specification while section 5 reports descriptive statistics for both types of households and compares older and younger siblings between them and across household types. Section 6 shows the main results while section 7 provides interpretations of the findings. Section 8 discusses possible mechanisms behind the results while section 9 provides a number of robustness checks. Finally, section 10 concludes.

2 Literature Review

The effects of adverse conditions during the early years materialize very soon in life, as gaps in skills arise very early between individuals and across socioeconomic groups (Cunha & Heckman 2007).[4] Moreover, gaps in development across the wealth distribution widen as children get older.[5] Inequalities at such an early stage is a concern as early childhood development has long-lasting consequences in adult success such as educational and labor market outcomes.[6] Hence, understanding what is behind the human capital production function has become a topic of interest in development economics. From a policy perspective, the question narrows to when and how to intervene, in order to ensure equal opportunities for children to reach their development potential.

There has been a significant progress in the knowledge of how skills are accumulated. Child development is multidimensional; skills can be categorized as cognitive-language, sensory-motor or socio-emotional skills. Following the theoretical model by Cunha & Heckman (2007), skills formation is a multistage technology, as it is a process that occurs at different stages of the child's life. Each of these abilities is constructed on top of previously developed skills. Moreover, every element interacts with each other and previous investments reinforce future ones. There are two important aspects worth highlighting about the human capital accumulation process. First, there are critical periods in which some skills are more easily acquired than at other stages, making the timing of investment a first order concern. Second, while genes play an important role in the process of skills formation, external factors such as parenting are fundamental during the first years of life (Shonkoff et al. 2012, Cunha et al. 2006).

Impact evaluations on home visit programs have provided good evidence of how parenting impacts child development. One of the most well-known evaluations took place in Jamaica, in which apart from nutritional supplements, visitors taught mothers to play with their children. Grantham-McGregor et al. (1991) find that stimula-

[4]Evidence can be found in contexts as diverse as India, Indonesia, Peru and Senegal (Fernald et al. 2012), Chile, Colombia, Ecuador, Nicaragua and Peru (Schady et al. 2015) and Philippines (Ghuman et al. 2005).

[5]This pattern has been found among children aged 6-42 months in Colombia (Rubio-Codina et al. 2015), aged 3-6 years old in rural Ecuador (Paxson & Schady 2007) and Bangladesh (Hamadani et al. 2014), and aged 7 months-5 years old in Madagascar (Fernald et al. 2011) to name a few.

[6]Key references include Currie & Almond (2011), Grantham-McGregor et al. (2007), Hoddinott et al. (2013), Black et al. (2017).

tion had a positive effect on children's development, and this effect lasted after 20 years (Gertler et al. 2014). Another evaluation in a psychosocial stimulation program in northeast Brazil also found positive results on cognitive and motor development (Eickmann et al. 2003). Likewise, in a randomized controlled trial in Colombia it was found that stimulation improved cognitive scores as well as receptive language (Attanasio et al. 2014). Attanasio et al. (2020) find that parental investment was the main driver of the improvement in skills. Moreover, they encounter that higher skilled mothers invest more in their children, once accounting for their children's skills level, highlighting the relevance of maternal characteristics. While these interventions have enlightened the role of parents in the development of children, the focus has been mainly on the mother as the main provider of stimulation. However, there are reasons to believe that inputs by the father could be complementary in the human capital production function, as shown by Macours et al. (2015). When testing the impact of a home-visiting parenting intervention in rural Nicaragua, they find that the effect on abilities is larger when both the mother and the father take part in the intervention. These findings suggest that the father may be more than the household's breadwinner. These results raise the question of complementarity of maternal and paternal inputs in skill formation, motivating the need to further investigate how fathers matter in early childhood development.

When it comes to critical periods, the medical literature has stressed the relative importance of the first 1,000 days of life for the development of the child (Barker 1990 and Victora et al. 2010). The economic literature has focused on the economic return to early interventions, arguing that investments during the early years is often more cost effective than later remediation interventions (Carneiro & Heckman 2003, Currie & Almond 2011), some emphasizing on the pregnancy period (Doyle et al. 2009). While still in their infancy, there have been some attempts to empirically identify the critical windows for early interventions in developing countries. Pathak & Macours (2017) analyse in India how policies that could potentially benefit human capital development can have long-lasting impacts in learning when children are exposed during the first 1,000 days versus at a later stage. Results show their educational outcomes are larger when children take advantage of those policies very early in life. Barham et al. (2013) compare the effect of a conditional cash transfer among Nicaraguan boys at age 10 exposed to the program before age 2 versus those exposed between age 2 and 5. They find short-term improvements in nutritional and health outcomes while in their long-term evaluation they encounter higher cognitive skills but equal anthropometric outcomes, consistent with a catch up in nutrition. This last finding provides evidence on the singularity of sensitive periods for each skill. Furthermore, it triggers the question on the mechanisms behind the results. More specifically, whether paternal and maternal inputs are equally relevant during this critical window.

This paper aims at contributing to a relatively unexplored literature in economics by studying the specific role of the father during the critical periods for human capital development. In order to isolate paternal inputs, my approach is to study the extreme case in which the father is absent from the household. Aiming at finding an adequate counterfactual, I compare the development across siblings in the same household exploiting the differential in timing of paternal absence. Moreover, I focus on four countries (Ethiopia, India, Peru and Vietnam) which not only provides evidence from the developing world but also provides external validity given the diversity on economic, social and cultural characteristics across countries. Below, I discuss related literature and how this paper contributes to each field.

The most evident literature studying the role of the father on the child's life is the one on migration and union dissolution. Within the migration literature, the focus is on disentangling the multiple effects that paternal departure may have on children's well-being. The ambiguity comes from the fact that departure may lead to an income effect –loss of income source versus remittances sent– but also to the so-called disruption effect. Paternal migration may impact children emotionally while lack of time investment could potentially jeopardize the cognitive and socio-emotional development of children. Moreover, it could alter resources allocation as a result of a change in decision making inside the household. Some studies explore the positive effect of remittances on educational investment (Cox Edwards & Ureta 2003, Calero et al. 2009, Acosta 2006). Others have tried to disentangle the income versus disruption effect of parental absence, finding that remittances can eventually compensate the lack of parenting conditional on socio-economic factors (Amuedo-Dorantes et al. 2010), if at least one parent stays at home (Lu & Treiman 2011) while others encounter heterogeneous effects between boys and girls (Hu 2012). Some studies differentiate between international versus domestic migration (Adams Jr & Cuecuecha 2010, Antman 2012), while others examine how the effect varies depending on the length of migration episodes (Giannelli & Mangiavacchi 2010). The literature focusing on divorce has also provided useful insights on the ambiguity of paternal absence. In rural Malawi, Chae (2016) finds that parental divorce is associated with lower grade attainment and larger schooling gap among children. Thiombiano et al. (2013) find in Burkina Faso that children of divorced parents experienced higher risks of dying before age 5 while their probability of entering primary school is lower. In both areas, little is known about the effects of father's departure during the critical periods neither on skills measured in the early years. One relevant study is Macours & Vakis (2010), who find that seasonal migration of the father in Nicaragua has no significant effect on early cognitive development.

A literature closer to the role of the father at an early age is the one on paternity leave. A few papers have examined the impact of taking paternal leave on fathers' participation in childcare; such as Nepomnyaschy & Waldfogel (2007) in the United States, Tanaka & Waldfogel (2007) in the UK, Ekberg et al. (2013) and Haas & Hwang (2008) in Sweden, Rege & Solli (2013) in Norway. Nevertheless, there is limited evidence on the implications of paternity leave on children's outcomes. One of the exceptions is the study by Cools et al. (2015) in Norway, where they take advantage of the paternal quota in 1993 and find that children's school performance at age 16 increases for children whose fathers are more educated than the mother. Again, effects are measured when children are no longer in their early years. Moreover, most of this literature has been focused on developed countries.

One of the strengths of this paper is the suitability of the data used to properly overcome previous papers' limitations, i.e. early episodes of paternal absence and ability outcomes measured early in life. Fortunately, Young Lives project provides a unique dataset in that sense, as it is a longitudinal study that follows children from an early age in the four mentioned countries. Young Lives reports detailed information on children and household characteristics. First, it includes information on the frequency in which children see their parents, which can be used as a proxy for paternal absence. Second, it provides measures on cognitive and socio-emotional outcomes from a young age.

Other researchers have taken advantage of the richness of Young Lives panel data to investigate similar interesting questions. Focusing on the role of the father, Dearden et al. (2013) assess the impact of father presence on Peruvian children's growth, making the distinction between absence during infancy, childhood and both. Using a mixed linear regression model, they find a negative effect on height when fathers are absence in both periods versus when he is present. On the other hand, children experiencing paternal absence during infancy (but not at age 5) performed better than those with a present father in both periods. Moreover, Nguyen & Vu (2016) assess the effect of temporary migration of parents on children's outcomes when aged 5 and 8 in Vietnam. With a child fixed-effect approach, they find that parental absence has an effect on children's time use. More specifically, children spend less time on home study, being the effect higher when the mother is the one absent (versus the father). In addition, maternal absence is linked to more time spent on housework.

Nguyen (2016) performs a similar analysis in Ethiopia, India, Peru and Vietnam, and finds that parental migration helps households increase per capita consumption but does not improve nutritional and cognitive outcomes of children. Parental migration decreases health outcomes in India, Peru and Vietnam and reduces cognitive skills in India and Vietnam; and the effect accentuates the longer the parental departure is. In Ethiopia the effect is not statistically significant.

Using child-fixed effects, as the latter studies do, solves the differences in time-invariant characteristics among households. However, just following the same child over time does not allow to study the consequences of paternal absence when very young. A more suitable counterfactual of the child is needed to eventually be able to find a causal relationship between absence at an early age and child's outcomes. The use of siblings has been a common practice to deal with this challenge (Behrman & Rosenzweig 2004, Black et al. 2005). Antman (2012) combines both the study of paternal migration in Mexico and the use of siblings to understand the impact of the father's absence on schooling outcomes. By assuming educational outcomes are not affected when children are at least 20 years old, she compares older versus younger siblings who may still be in school. Using variation in siblings' ages at the time of parental departure is an interesting approach to overcome selection problem. Indeed, households with an absent father might differ from those with no absence both in observable and unobservable characteristics. Therefore, a sibling seems to be a good counterfactual of what would have happened to the child under a different context.

Fortunately, Young Lives includes a survey on siblings, which has already been exploited in several studies. Sanchez et al. (2016) use the paired-sibling sample in Peru to evaluate the effect of conditional cash transfer (Juntos) implemented in 2005 on nutritional and cognitive outcomes. They take advantage of the difference in the timing in which siblings benefited from the program. Similarly, Outes-Leon et al. (2011) use a sibling-difference specification for Peruvian pre-school children in order to assess the link between nutritional outcomes and cognition, having as an instrument household shocks and food price changes. They compare older and younger siblings aged approximately 4-6, but measured at different points in time, 2006 and 2009 respectively.

In this paper, I rely on the availability of development outcomes of siblings at different stages in life and take advantage of the timing in which paternal absence occurs. I control for the differential in development trends between siblings who had not experienced the departure at all. In this way, I am able to find a causal effect of the absence of the father on a child development when it occurs during the critical window versus at a later stage. In the following sections I explain the data used and the empirical methodology in more detail.

3 Data

The data comes from Young Lives[7], an international study of childhood poverty in Ethiopia, India (Andhra Pradesh and Telangana), Peru and Vietnam. It follows two cohorts of 12,000 children in total for over 15 years. The first cohort was born in 1994-95 while the second was born in 2001-02. This paper focuses on the younger cohort, for which five waves were used: 2002 (the baseline), 2006, 2009, 2013 and 2016. Attrition rates between the first and fifth round are relatively low, around 4.5% for the four countries. In terms of sampling approach, around 20 sentinel sites were selected non-randomly in each country and poor areas were over-sampled. The original sample comprises of around 2,000 children in each country, who were randomly selected based on their age from each of these communities. From now on, these children will be referred to as the *index children* or *older sibling*.

[7]Huttly & Jones (2014), Boyden (2014), Boyden (2011) Boyden et al. (2016) and Sanchez et al. (2018).

Young Lives longitudinal study is a unique dataset, as it provides detailed information on households and community characteristics and key questions about parents which I use to proxy paternal absence. My definition of absent father is based on the information collected on the frequency by which the index child sees her/his father as well as if the father lives at home. Additionally, the survey is particularly suitable for my analysis due to the inclusion of development outcomes of pairs of siblings in each household. While originally only the index child was tracked, information on the closest-in-age sibling started to be included after the third wave. The sibling questionnaire comprises those who were at least 24 or 36 months old (depending on the country), were able to stand and were in the community at the time of the survey. I will use data on all siblings who were born between 2003 and 2005.

The identification strategy relies on the availability of development outcomes for both the older and the younger sibling when 5, 8 and 12 years old plus information on father's involvement; allowing me to compare pairs of children from the same household that experienced absence at different moments in time. The following section explains how these variables are used to capture the causal relationship between paternal input in the early years and childhood development. But first, I provide more details on how paternal absence is defined, how the sample of siblings is selected as well as the construction of development outcomes for both children.

Paternal absence

From the question "How often does the child's father see the child?", an absent father is defined as a parent who does not see his child on a daily basis or if he does, the father does not live at home.[8] In terms of residency of the father, round one includes the following question: "Where does child's biological father live?". For the following rounds, IDs were matched in order to check their presence in the family roster and whether the individual still lived at home. It is important to note that in some rounds it is specified the fact that the father is the biological one. In round two and three, an extra question is included, where it is explicitly asked whether the paternal figure is the biological father. As it seems too restrictive, I do not constraint by the fact that the father is the biological one in the main specifications as there are only a few cases in which the father was not.[9]

Siblings sample

Siblings were incorporated in the survey from round 3 (2009) onward, including information such as their birthdate, anthropometric measures, cognitive and socio-emotional tests. In order to find a suitable counterfactual of the index child, I construct the appropriate sample by matching pairs of siblings who were at a similar stage in life one round apart (i.e. the older sibling should be the same age in 2006 as the younger sibling in 2009).

Unfortunately there is a trade-off between how close siblings are in age and sample size. Aiming at not being too restrictive, I impose that the younger sibling should had been born between 2003 and 2005 (between round 1 and round 2). On the downside, this leads to a wider distribution in terms of age among younger siblings than among index children.

[8] The question varies among rounds. In round 1 the exact question is the following: "In the last 6 months how often has the child's biological father seen him/her?", in the second and third round: "How often does the father see the child?". In the fourth and fifth round the following question is only asked when the biological father is not in the household roster: "How often does the father see the child?".

[9] If I calculate absence restricting to being a biological father, I would have 3 more observations when absence occurs in 2006.

First, I include those siblings already born in 2006, based on their presence in the household roster in round 2. Furthermore, I exclude those already born in the first round (2002), based on their appearance in the family roster or on the age they would have been in that round.[10] I calculate age from birthdates when available/reliable, otherwise I use reported age. I delete those observations where gender does not match between round 2 and 3. I also restrict to those who are biological brothers/sisters in order to make sure they share the same paternal figure. I end up with 2,565 younger siblings. For further information on how age in months was calculated, see appendix A.1. Table A1 exhibits the distribution of rounded ages between the index children and the younger siblings.

Development outcomes

In this paper, the outcomes under analysis are on nutrition, such as weight-for-age and height-for age, cognition- Peabody Picture Vocabulary test, and socio-emotional skills, agency and pride index.

Nutritional outcomes are calculated from information on weight, height, gender and age in months. Anthropometric z-scores are computed using the 2006 WHO child growth standards for children between 0 and 5 years old and 2007 WHO child growth standards for children aged 5 to 19. For weight-for-age, growth standards are only available up to age 10. For the index children, age in months is available in the dataset while for the younger siblings it is calculated based on birthdates if not included in the dataset (see appendix A.1 for more details). Z-scores are considered missing if values are implausible (based on the WHO cutoffs).

The Peabody Picture Vocabulary Test (PPVT) is a widely used cognitive test that attempts to measure children's vocabulary by showing them a set of four pictures simultaneously and requesting them to select the picture that best represents the word mentioned by the examiner. Whereas for the index child data is available for the four countries, it is not the case for the younger sibling as the test was only administrated in Ethiopia, Peru and Vietnam (provided he/she was at least four years old in 2009). Therefore, India is excluded for this outcome. In Ethiopia and Vietnam the version used was the PPVT-III, while in Peru the Spanish version of the PPVT-R was provided. These two versions differ in aspects such as the number of items and the definition of the basal and ceiling rules, meaning that the computation of the raw score is not done equally. Moreover, raw scores are comparable only among children of the same age, which makes this score less suitable as the range of ages between pair of siblings widely varies. For these reasons, I consider that a proper standardization of the test needs to be done. Using external z-scores may not be the best option as external norms might not be appropriated when using data from developing countries. Hence, I use internally standardized scores by performing a simple standardization where the mean and standard deviation are calculated per country pooling older and younger siblings at all ages. As the variance in age among siblings is higher than among index children, and because adjusting for age is a common practice, I will control for a polynomial of age in the regressions.

The survey includes a group of psychosocial measures, pride and self-esteem (based on Rosenberg 1965) and agency/self-efficacy (based on Rotter 1966 and Bandura 1993). These items are only included in round 4 and 5 for the sibling. All questions are answered on a Likert scale (from 1 to 5, 5 being strongly agree). For each index, three subindices were constructed taking into consideration the fact that some questions were only asked to those children at school and those that worked. For the agency indices, some questions were asked negatively, which could eventually arise the problem of acquiescence bias, which refers to the tendency of the respondent

[10] Both conditions were imposed as some siblings were not in the household roster in 2002 but in fact were already born based on their calculated age. Therefore, siblings were kept in the sample as long as they were born at least one month after the first round survey was done.

to agree more than disagree. In order to correct for this bias, I follow the approach by Laajaj & Macours (2017) and calculate the acquiescence score of each individual by averaging the mean of all positive items and the mean of the negative items (before reversing) and then average both. After subtracting 3, I deduct the acquiescence score to each question. All items were standardized using the mean and standard deviation per country. Each question in each subindex was weighted using factor analysis technique. For more details on items included in each index, see appendix A.2.

Once again, measures on anthropometrics and cognitive skills for the younger sibling are available in the last three waves, while information on socio-emotional skills is available in the fourth and fifth waves.

4 Identification Strategy and econometric specification

4.1 Identification Strategy

The aim of this paper is to estimate the impact of an absent father during the critical periods on early childhood development. When estimating the causal impact of the absence of the father on early childhood development, researchers face the endogeneity related to the departure of the father and children's performance. An appealing approach is to use siblings as counterfactuals, given that they are a good approximation of what would have happened to the child under different external conditions. Under this rationale, I evaluate the differential effect of paternal absence in the early yearstaking abvantage of the timing of this absence and the age of the older and the younger sibling at which this event took place. By this way, I try to mimic the older child's development path in a scenario in which absence of the father had occurred in the early years instead of when aged 5. In order to approximate my analysis as much as possible to this ideal scenario, I compare siblings when they are around the same age.

The simple comparison of siblings in households that experienced the absence of the father would ignore two issues. First of all, parents might allocate investments differently across siblings as a result of preferences or due to the economic cycle of the household. Secondly, as siblings are measured at different points in time, certain common shocks (for instance, a macroeconomic crisis) could have affected development independently of the departure of the father. One way to address these issues is to control for the eventual difference in development between siblings by including pairs of children from households with no paternal absence.

The identification strategy relies on a central assumption, which is that the potential gap in development between the older and younger sibling in households that experienced the absence of the father would have been the same as those with no absence if the father had not left. In other words, trends in nutritional, cognitive and socio-emotional outcomes between siblings would have been parallel between both type of households if the father would have been present. Section 9 provides support for this assumption. This assumption could be violated if there is a shock that has an differential impact on children's well-being and at the same time leads to the absence of the father. For instance, families could decide to send the father away as a result of the sickness of one of the children. If that was the case, I would be concluding that paternal absence has a negative effect on health when in fact departure was a response to the child's illness. For this reason, I will control for the fact the family experienced a negative shock as a robustness check. This will be explained in further detail in section 9.

Taking this assumption into consideration, I exploit Young Lives data available on the children and their younger siblings born in 2001-02 and 2003-05 respectively and the possibility that a pair of siblings had experienced different paternal involvement early in life. The panel data will comprise of pairs of siblings when approximately aged 4-6 (from now on 5 year-olds) and 7-9 (from now on 8 year-olds). Alternative specifications used throughout this paper are explained in section C, E and F.2. Below, the details on how I use this data for the identification strategy.

4.1.1 Paternal absence during the early years versus at age 5

As seen in Figure 1 the first group (treatment group) had experienced the absence of the father in 2006 but not in 2002. In those households, only the younger sibling is the one who faced the absence of his/her father when very young, but not the older sibling who did when aged 5. In the control group, neither the younger nor the older sibling experienced the absence of the father in 2006 nor in 2002. Note that there is only information if the father was present or absent each round but not on the date when the departure took place. The treatment group is composed of 153 households while the control group comprises 2,212 households (see table B3 for details on sample size per country).

There is no restriction on the presence/absence of the father in 2009 and onward. This is relevant as it could be the case that households with no absence in 2006 do experience the departure of the father afterward, in which case I would be comparing pairs of siblings experiencing the absence of the father at different moments in time. On the other hand, comparing households that have the father gone at some point may be more similar to the treated families, making it more accurate as a control group. As a robustness check, I will include alternative controls restricting presence of fathers after 2006 (Section 9).

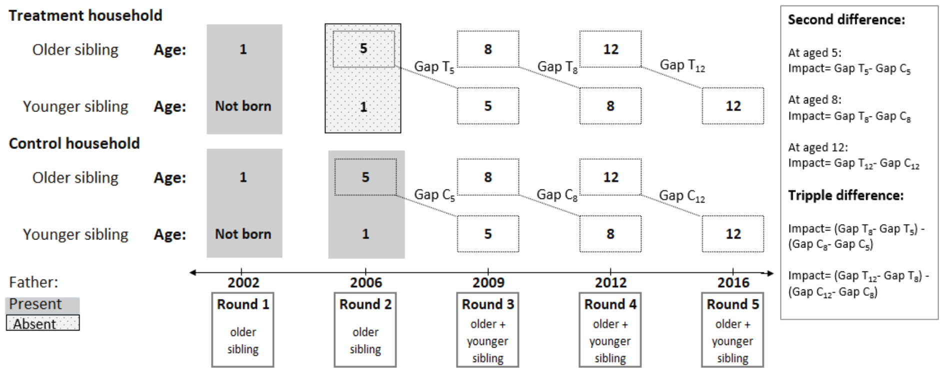
In this exercise I take the first difference in nutritional and cognitive outcomes between pairs of siblings when around 5 years old (for the older sibling in 2006 and for the younger one in 2009) from treated households (absence in 2006) and the same difference between pairs of siblings in households in the control group. I perform a similar analysis when both siblings are 8 and 12 years old in order to have some insight on long terms effects. For socio-emotional outcomes, the analysis is done when siblings are around 8 and 12 years old.

Examining medium and long term effects is relevant as these two effects could differ for several reasons. For instance, the effect of the absence of the father that occurs when the younger sibling is very young, could potentially vanish, persist or increase over time (independently if the father comes back or not). In the case of the older sibling, if the departure of the father was very recent in 2006, measuring her/his outcomes when 5 years old may be premature to actually capture any kind of effect and would be likely noticeable at a later stage.

If there is evidence of medium and long term effects, it is interesting to understand the trends across time. Hence, as a secondary exercise, I study how paternal departure affects the evolution of these outcomes by performing a difference in the change when children are 5 and 8 years old, between pairs of siblings, among households experiencing the absence of the father and those that do not. Similarly, I study the change when children are 8 and 12 years old.

Before proceeding, a few data limitations for my paper need to be mentioned. Given the sample size of treated households, one important concern is that this could compromise power and potentially increase the probability of type II error. For this reason, I restrict the sample the least possible by including siblings who were around– but not exactly– 5 years old in 2009. Moreover, countries are aggregated in order to gain the most feasible power. However, pooling countries together requires making the assumption that effects of the absence of the father are not heterogeneous and trends among siblings are similar across countries. I test for differences between countries.

Figure 1: Absence in 2006: timing of absence and measurement of children's outcomes



Notes: This figure shows the identification strategy for the exercise. There are two types of households: (1) treated households, the father was present in the first round but was absent in 2006 (2) control households, the father was present in the first two rounds. In each household there is information on nutritional, cognitive and socio-emotional outcomes for a pair of siblings, a younger and an older sibling. Therefore, the sample includes those households with at least two siblings. Information on outcomes is only available for the younger sibling from 2009 onward. The main analysis consists of calculating the difference (or gap) between the younger and older sibling in the treatment group when 5, 8 and 12 years old (Gap T_4, T_8 and T_12 respectively) minus the gap between siblings in the control group (Gap C_4, C_8 and C_{12} respectively). As a ancillary exercise, I also calculate the difference in the evolution of the gap between siblings in the treatment group versus the control group when they are 8 versus 5 ((Gap T_8- Gap T_4)-(Gap T_12- Gap T_8)) and 12 versus 8 ((Gap T_8- Gap T_4)-(Gap T_12- Gap T_8)).

A second limitation is that, while there is information on the absence of the father at specific points in time, it is not known whether fathers were absent in-between rounds. As long as there is no clear pattern among the control versus treated households in-between waves, this should not harm the credibility of my results. Moreover, there is no information on the duration of absence as there is no precise data of the departure date. This could be an issue if the length of absence is not long enough to have an effect on development outcomes. Estimating long term effects allows me to check whether paternal absence becomes more salient after a while among those in which fathers had just left in the studied rounds.

Lastly, the definition of absence is based on the frequency the child sees his/her father and if he lives in the household. However, there is no information on the reason the father has left (except if they are dead). Absence can be due to permanent or seasonal migration, or can be the result of abandonment, divorce or death. Restricting future presence gives some sense of the most plausible reason of departure. For instance, if the father is back in the next round then it is probable that he temporarily migrated. In section section 8.4 I will study cases in which the father comes back the following round versus those that do not.

4.2 Econometric Specification

I propose the following baseline econometric specification, represented by equation (1):

$$\text{y}_{ih} = \alpha + \gamma \text{younger sibling}_{ih} + \delta(\text{HH w/absence in } 2006_h \times \text{younger sibling}_{ih}) + \text{HH}_h + \phi X_{ih} + \epsilon_{ih} \quad (1)$$

Where y_{ih} are nutritional (weight-for-age and height-for-age) and cognitive outcomes (standardized PPVT) when paired-siblings are around 5, 8 and 12 years old, while it represents socio-emotional outcomes when children are 8 and 12 years old. Variable $younger sibling_{ih}$ is equal to 1 if the outcome belongs to the younger sibling and 0 if the outcome belongs to the older sibling. $HHw/absence\ in\ 2006_h$ is equal to

1 if the household experiences absence in 2006 (but not in 2002) and 0 if there was no episode of absence in these two rounds. The coefficient of interest is δ. HH_h is the household fixed effect. X_{ih} includes age, gender and birth order of each child. It also includes the interaction between each control variable and country dummies. When the dependent variable is the cognitive test measure, I include an age polynomial as control. Standard errors are clustered at the sentinel site level.

5 Descriptive statistics

Table B4 reports information on the characteristics of older siblings, mothers, fathers and their households at baseline (i.e. 2002) by household type. Questions on whether the mother/father takes care of the child and if they are a financial support are referred to the older sibling (index child). On average, older siblings in both groups are similar in terms of age, birth order, reported birth weight, reported number of prenatal visits by mothers and whether the child attends regularly a *creche*. There are more girls in households where no absence occurred, significantly different at the 10% level. In terms of mothers' characteristics, they are slightly older in households experiencing paternal absence in 2006 while they do not seem to differ much in terms of education. In both types of households mothers regularly take care of the child (96%) and 63% are said to provide financial support to the older sibling. Fathers differ in terms of education as those who were absent in 2006 are on average more educated than those who were present. Fathers seem on average to equally take care of the older sibling and to provide financial support at baseline. Household size does not statistically differ across type of households neither the percentage of families who report to live in the rural area. Moreover, a statistically similar percentage of household stated that a member owns the house. However, households in the treatment group seem to be on average richer when looking at the housing quality index and wealth index. Furthermore, treated households are more likely to have experienced a negative shock at baseline than households in the control (62% versus 50%) and this difference is statistically significant at the 1% level.

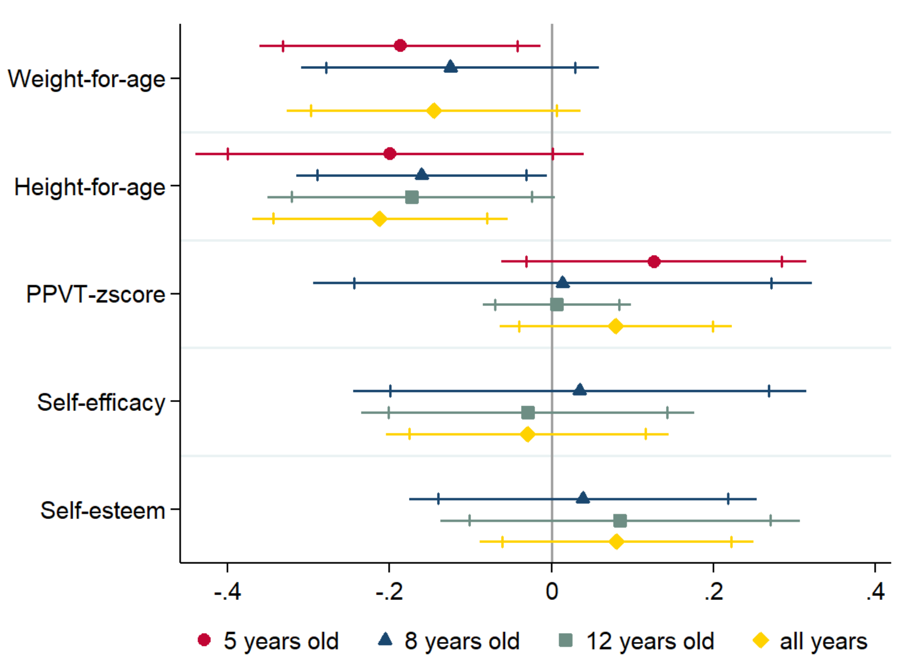
Table B5 exhibits descriptive statistics of the outcomes of siblings (both older and younger ones) when 5 years old in households with and without an absent father in 2006. In general terms, older siblings in households that experience absence when aged 5 outperform on average older siblings in households with no absence in terms of nutritional outcomes and these differences are statistically significant at the 1% level for height for age and at the 5% level for weight for age. However, their cognitive test scores are not statistically different from each other. The percentage of girls in households with an absent father is lower than the one in households not experiencing absence in the selected rounds. On the contrary, younger siblings are not statistically different from each other across type of households in terms of nutrition (standardized measures) and cognition, except for age, as children in households with an absent father early in life are slightly older than those in households with a present father. When looking at the gap within household types, column 13 shows that older siblings are doing better in terms of weight and height for age compared to their younger counterparts in households with the absence episode. However, younger children are outperforming the older ones in terms of cognition. Those differences are statistically significant. In households with no absence, older siblings outperform younger ones in terms of weight and height but when age differentials are taking into consideration (i.e. standardized values), the difference is no longer significant. As in treated households, the gap in cognition is positive and statistically significant between the younger with respect to the older sibling. It is worth noting that the gap in nutritional (standardized) outcomes and in cognitive skills is more accentuated for the pair of siblings experiencing the absence of the father at different moments in their lives than for those that do not.

6 Results

6.1 Pooled results

Figure 2 presents the main aggregated results for the estimate of interest, the interaction between being a younger sibling and belonging to a household that experienced the absence of the father in 2006. The figure includes estimates when children are 5, 8 and 12, as well as the average of all available years together[11] for nutritional, cognitive and socio-emotional outcomes. The missing results of weight-for-age at 12 years old is due to the lack of benchmark for children over 10 years old to calculate the standardized score. Self-efficacy and self-esteem are available for the last two rounds and correspond to a shorter index which excludes questions related to school and work.[12] Table 1 shows the corresponding regressions for each period and the average of all ages together. Once again, both socio-emotional indexes correspond to the shorter version. For results using alternative definitions (including questions on schooling and/or work), refer to table C6.

Figure 2: Coefficient of interest: nutritional, cognitive and socio-emotional gap when 5,8 and 12 years old and all years together



Notes: This figure reports the estimates of interest (hh with an absent father in 2006 and outcomes for the younger sibling) obtained from the estimation of equation (1), when siblings are 5, 8 and 12 years old and all years together. The figure includes 90% and 95% confidence intervals. Estimates represent the gap in the corresponding outcome between a child and his/her younger sibling in households that experienced the absence of the father in 2006 versus the gap in household that did not experienced his absence that year. First two dependent variables correspond to weight-for-age and z-score and height-for-age z-score. Both anthropometric z-scores are based on the 2006 WHO child growth standards for children below 5 and 2007 WHO child growth standards for children aged 5-19 years (except for weight-for-age for children older than 10). The third dependent variable corresponds to the standardized score of the Peabody Picture Vocabulary Test (PPVT) at the country level. Last two dependent variables are self-efficacy and self-esteem, which are available for the last two rounds and correspond to a shorter index which excludes questions related to school and work. For cognitive outcomes, India is excluded due to lack of data for younger siblings. All regressions include a household fixed effect and are controlled for age, gender and birth order. Interactions between each control and each country are included.

[11] Refer to section C for details on the specification strategy

[12] This aims at restricting the sample the least possible, given that for those children that do not go to school or do not work, related questions appear as missing.

I begin by looking at anthropometric outcomes, weight and height for age. In households where the father was present, the nutritional development of younger siblings is statistically similar to that of their older siblings when aged 5 (see first row, column 1 and 2 of table 1. When looking at the coefficient of interest, the interaction between being the younger sibling and belonging to a household with an absent father in 2006, I find it is negative for both weight and height for age. This implies that, while pairs of siblings in households with no absent father perform similarly, the gap between the older and younger sibling become negative and significant in households with an absent father (for weight-for-age)[13]. These results suggest that the absence of the father during the first years of life affects negatively (or more negatively) the nutritional development of children compared with the case where absence occurs at age 5.

The third panel of figure 2 reports cognitive skills, proxied by the Peabody Picture Vocabulary Test (standardized). The sample excludes India, leading to a lower sample size of households with respect to nutritional outcomes. First row of column 3 in table 1 shows that in households with a present father, the younger siblings outperform on average their older counterparts in terms of cognition, and this difference is statistically significant. When looking at the coefficient of interest, I find that this positive gap increases in households in which the father was absent. While not statistically significant, this is a striking result as it shows a differential effect with respect to nutritional outcomes. Moreover, lack of power could be hiding a positive impact (in relative terms) of the absence of the father on cognition. On table C9 in the appendix, I perform a sensitivity analysis, looking at results using the raw score, using an alternative standardized score (conditional mean), excluding outliers (>4 sd) and restricting to the same sample in the three rounds). General conclusions of a non-negative coefficient for 5 year old children is consistent across these specifications.

To summarize, I find that in the short run (when children are 5 years old), there is a negative gap on nutritional development between younger and older sibling in households with an absent father. On the contrary, I find a null effect on cognitive outcomes when measured at that same age.

Up to this point, I have been comparing siblings that experienced the absence in the early years (younger siblings) and those that experienced it when 5 years old in treated households versus pairs of siblings in households from the control group. Since outcomes are contrasted when pairs of siblings are aged 5, the older sibling is being measured at the same wave at which the father is reported to be absent in treated households. It might be the case that departure has been so recent that older siblings had not yet faced the consequences of paternal absence. Therefore, looking at the evolution of the gap seems relevant not only to better understand the medium and long term effect of the absence but also in case the effect on the older child becomes apparent after a while.

When comparing the development of siblings in the medium and long run, the estimate for weight-for-age still suggests a negative effect when they are measured at 8 years old (while no longer significant at the 10% level). In line with this result, estimates on height-for-age show a persistent relative under-performance between the younger and the older sibling in the medium and long run, being statistically significant at the 5% level when 8 and at the 10% level when 12 years old. The richness of YLS allows me to look at two measures of socio-emotional skills, self-efficacy and self-esteem for both siblings when 8 and 12 years old. From figure 2 I conclude there is no evidence of a differential response of children based on the age they experienced the absence of the father.

So far, the results presented correspond to the specification in which countries are

[13]While it is significant at the 5% only for weight-for-age, the p-value for height-for-age is around 0.102, close to be significant at the 10% level.

pooled together. However, I could expect heterogeneities to arise across countries. This is particular relevant as I am including different cultural contexts which may broadly vary in terms of family structures and family interactions and investments.

Table 1: Nutritional, cognitive and socio-emotional outcomes when 5, 8, 12 years old and all years together

VARIABLES	(1) Weight for age	(2) Height for age	(3) PPVT	(4) Self efficacy	(5) Self esteem
5 years old					
Younger sibling	0.113	0.407	0.479**		
	(0.255)	(0.341)	(0.202)		
HH w/absence in 2006 * younger sibling	-0.187**	-0.199	0.126		
	(0.0870)	(0.120)	(0.0942)		
Observations Control HH	2145	2122	1219		
Mean Control (older siblings, no absence)	-1.34	-1.64	-.89		
Observations Treatment HH	149	148	95		
Mean Treatment (older siblings, absence)	-1.1	-1.37	-.97		
8 years old					
Younger sibling	0.645***	0.853***	-1.659***	-0.482*	-0.643***
	(0.170)	(0.182)	(0.388)	(0.244)	(0.243)
HH w/absence in 2006 * younger sibling	-0.125	-0.160**	0.0138	0.0346	0.0388
	(0.0924)	(0.0775)	(0.154)	(0.140)	(0.108)
Observations Control HH	1700	2132	1347	1924	2036
Mean Control (older siblings, no absence)	-1.45	-1.38	.78	-.07	.13
Observations Treatment HH	106	149	109	138	147
Mean Treatment (older siblings, absence)	-1.19	-1.21	.87	-.09	.05
12 years old					
Younger sibling		0.147	-0.228	0.443	0.0857
		(0.220)	(0.153)	(0.266)	(0.159)
HH w/absence in 2006 * younger sibling		-0.173*	0.00669	-0.0290	0.0845
		(0.0890)	(0.0457)	(0.103)	(0.111)
Observations Control HH		2117	1415	2110	2127
Mean Control (older siblings, no absence)		-1.44	.28	.08	-.12
Observations Treatment HH		148	110	147	148
Mean Treatment (older siblings, absence)		-1.21	.32	.1	-.14
5, 8 and 12 years old together					
Younger sibling	0.231***	0.423*	-0.187	0.0696	-0.162
	(0.0529)	(0.228)	(0.144)	(0.187)	(0.117)
HH w/absence in 2006 * younger sibling	-0.0632*	-0.212***	0.0795	-0.0296	0.0802
	(0.0362)	(0.0792)	(0.0715)	(0.0876)	(0.0850)
Observations Control HH	6333	6162	3318	3758	3990
Observations Treatment HH	444	435	267	264	284
Includes India	Yes	Yes	No	Yes	Yes
Age polynomial	No	No	Yes	No	No

Notes: Standard errors in parentheses, clustered at the sentinel site level. *** p<0.01, ** p<0.05, * p<0.1. Each regression includes a household fixed effect and presents the gap in the corresponding outcome between a child and his/her younger sibling in households that experienced the absence of the father in 2006 versus the gap in household that did not experienced his absence that year. All outcomes are measured when both siblings are around 5 (first panel), 8 (second panel), 12 (third panel) years old and all years together (fourth panel). Variable *younger sibling* is equal to 1 if the individual is the younger sibling and 0 if the individual is the relatively older child. Variable *hh with absence in 2006* is equal to 1 if the pair of siblings under comparison belong to a household which experienced the absence in 2006 but not in 2002, and 0 if the father was present in both rounds. Dependent variables in column 1 and 2 are weight-for-age z-score and height-for-age z-score. Both anthropometric z-scores are based on the 2006 WHO child growth standards for children below 5 and 2007 WHO child growth standards for children aged 5-19 years (except for weight-for-age for children older than 10). Dependent variable in column 3 is the standardized score of the Peabody Picture Vocabulary Test (PPVT) at the country level. For this outcome, India is excluded due to lack of data for younger siblings. Last two columns correspond to self-efficacy and self-esteem, which are available for the last two rounds and correspond to a shorter index which excludes questions related to school and work. All regressions are controlled for age, gender and birth order. Interactions between each control and each country are included.

6.2 Heterogeneous effects by country

Table 2 presents the same analysis as before, but including country interactions for both the dummy variable corresponding to the younger sibling and the coefficient of interest (being a younger sibling in a household with an absent father in 2006). I look at nutrition and cognition at 5 years old, but not at socio-emotional outcomes given that these measures are not available when children are this age. The table includes in brackets under each coefficient the p-values corresponding to the coefficient of the individual regression per country. It also provides the jointly significant test performed to assess whether country interactions are statistically different from each other.

Table 2: Nutritional and cognitive outcomes when 5 years old between younger and older siblings among households with and without and without absence in 2006 with country interactions

VARIABLES	(1) Weight-for-age	(2) Height-for-age	(3) PPVT Z-Score
HH w/absence in 2006 * younger sibling (Ethiopia)	-0.191	-0.0480	-0.0235
	(0.128)	(0.244)	(0.0925)
p-value individual regression	[0.160]	[0.849]	[0.805]
HH w/absence in 2006 * younger sibling * India	-0.248	-0.142	
	(0.188)	(0.317)	
p-value individual regression	[0.006]	[0.369]	
HH w/absence in 2006 * younger sibling * Peru	0.103	-0.378	0.231*
	(0.225)	(0.289)	(0.137)
p-value individual regression	[0.646]	[0.014]	[0.058]
HH w/absence in 2006 * younger sibling * Vietnam	0.177	-0.213	0.459*
	(0.251)	(0.323)	(0.271)
p-value individual regression	[0.950]	[0.245]	[0.111]
Number of HH	2,294	2,270	1,314
Observations control HH	2,145	2,122	1,219
Mean control (older sibling, no absence)	-1.34	-1.64	-.89
Observations treatment HH	149	148	95
Mean treatment (older sibling, absence)	-1.1	-1.37	-.97
P-value HH w/absence in 2006 * younger sibling	.272	.577	.107
Includes India	Yes	Yes	No
Age polynomial	No	No	Yes

Notes: Standard errors in parentheses, clustered at the sentinel site level. *** p<0.01, ** p<0.05, * p<0.1. Each regression includes a household fixed effect and presents, by country, the gap in the corresponding outcome between a child and his/her younger sibling in households that experienced the absence of the father in 2006 versus the gap in household that did not experienced his absence that year. All outcomes are measured when both siblings are around 5 years old. Dependent variables in column 1 and 2 are weight-for-age z-score and height-for-age z-score. Dependent variable in column 3 correspond to the standardized score of the Peabody Picture Vocabulary Test (PPVT) at the country level. For this outcome, India is excluded due to lack of data for younger siblings. All regressions are controlled for age, gender and birth order. Interactions between each control and each country are included. The table includes in brackets under each coefficient the p-values corresponding to the coefficient of the individual regression per country.

For nutritional outcomes, estimates of the interaction terms are not jointly significant. Hence, I find suggestive evidence that the effect found when pooling the four countries is representative of each one of them. In terms of magnitudes, India seems to be pushing the negative coefficient the most in terms of weight for age, while for height for age Peru is the country showing the most negative coefficient.

As for cognition, I observe that the coefficient goes close to zero for Ethiopia but this is not the case for Peru and Vietnam. In fact, estimates show a positive gap among pairs of siblings when the father left compared to those with no absence. Furthermore, the coefficient is statistically significant at the 10% level for Peru while close to be significant for Vietnam (p-value= 0.111).

Turning now to socio-emotional outcomes, table C8 present the desegregated results by country for self-esteem and self-efficacy when children are 8 and 12. I also include alternative definitions for both measures, adding questions liked to schooling and working. In Ethiopia, results reveal a relatively over-performance in terms of self-efficacy by the younger sibling when 8 years old in households with an absent father. However, this differential vanishes when measured the following round, when aged 12 years old. On the contrary, there seems not to be a differential gap in the development of this skill for the rest of the countries. The only exception is India with a significant coefficient when aged 12. When it comes to self-esteem (or pride index), Indian younger siblings seem to under-perform older ones at age 8 when suffering the departure of their father earlier on but it is only statistically significant when including all questions (column 4, smaller sample size). On the contrary, estimates are positive and significant for Vietnam for column 1 and 2 when aged 12. This gap only shows in the medium run as the gap disappears when they are 12. Given the noise these measures may have and lack of a clear trend, from now on I will focus on results on nutrition and cognition.

6.3 Evolution over time

Table D10 shows the differential in trends between sibling across household types, pooling back the four countries together. Details on the econometric specification for this exercise are explained in section E. I focus on nutritional outcomes given that pooled results are consistent with the individual results by country. First two columns show the evolution of the gap when pairs of siblings were 5 versus 8 years old. The last two columns show that evolution when siblings were 8 versus 12 years old. From table D10 I conclude that there is a persistent negative effect of the absence of the father on height-for-age, but these effects are no cumulative, as it does not increase nor decrease in time.

The magnitudes of the estimates do not seem to be statistically different in the short and long term. Hence, I pool all periods to gain power and compare the average gap between the younger and the older sibling for the three periods- when 5, 8 and 12 years old. Table 3 provides more precise estimates, showing that the relative deterioration in terms of weight and height-for-age are statistically significant. Except for short-term analysis, this will be the preferred specification given the relative advantage on power.

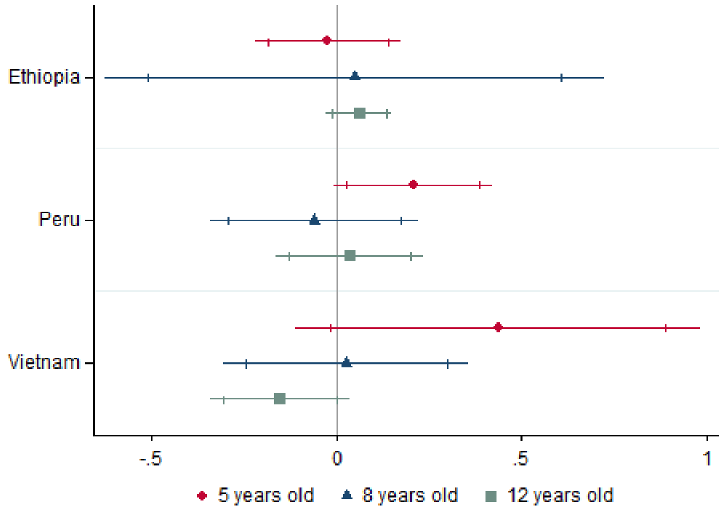
Table 3: Nutritional and cognitive outcomes (average when 5 and 8 and when 8 and 12 years old) between younger and older sibling among households with and without absence in 2006

VARIABLES	(1) Weight-for-age^{+}	(2) Height-for-age	(3) PPVT Z-Score
Younger sibling	0.200	0.423*	-0.187
	(0.140)	(0.228)	(0.144)
HH w/absence in 2006 * younger sibling	-0.145	-0.212***	0.0795
	(0.0910)	(0.0792)	(0.0715)
Observations Control HH	3213	6162	3318
Mean Control (older sibling, no absence)	-1.42	-1.48	.07
Observations Treatment HH	198	435	267
Mean Treatment (older sibling, absence)	-1.19	-1.26	.06
Includes India	Yes	Yes	No
Age polynomial	No	No	Yes

Notes: $^{+}$Excludes period when children are 12 years old.
Standard errors in parentheses, clustered at the sentinel site level. *** p<0.01, ** p<0.05, * p<0.1.Outcomes at 5, 8 and 12 years old were pooled together, and fixed effects for each wave were included (*8 yrs old* for the second period and *12 yrs old* for the third period). Variable *younger sibling* is equal to 1 if the individual is the younger sibling and 0 if the individual is the older child. Variable *hh with absence in 2006* is equal to 1 if the pair of siblings under comparison belong to a household which experienced the absence in 2006 but not in 2002, and 0 if the father was present in the first two rounds. Dependent variables in columns 1,2 and 3 are weight-for-age z-score (period 1 and 2 only), height-for-age z-score and PPVT z-score. Dependent variable in column 3 corresponds to the standardized Peabody Picture Vocabulary Test (PPVT) at the country level. For cognitive outcomes, India is excluded due to lack of data for younger siblings. All regressions are controlled for age, gender and birth order and whether the family experienced a negative shock in 2002. Interactions between each control and each country are included.

Figure 3 exhibits estimates for standardized cognitive measures by country. As previously shown, pooling the four countries does not seem to be pertinent given the heterogeneity found across Ethiopia, Peru and Vietnam. In the figure I observe that the null gap in Ethiopia remains statistically unchanged in the long run between pairs of siblings. In Peru, the positive and statistically significant gap at 5 years old vanishes over time. In Vietnam, the positive gap goes closer to zero and then becomes negative. While none of these coefficients are statistically significant for the latter country, this could suggest a reversal in the relative cognitive development as a result of paternal absence as time goes by.

Figure 3: Estimates when siblings are 5, 8 and 12 years old for cognitive outcomes per country



Notes: This figure reports the estimates of interest (hh with an absent father in 2006 and outcomes for the younger sibling) obtained from the estimation of equation (1), when siblings are 5, 8 and 12 years old for each country. The figure includes 90% and 95% confidence intervals. Estimates represent the gap in the corresponding outcome between a child and his/her younger sibling in households that experienced the absence of the father in 2006 versus the gap in household that did not experienced his absence that year. Dependent variables correspond to Peabody Picture Vocabulary Test (PPVT) standardized at the country level. For cognitive outcomes, India is excluded due to lack of data for younger siblings. All regressions include a household fixed effect and are controlled for age, gender and birth order.

7 Discussions

From the previous results I conclude that having an absent father very early in life has a differential effect than experiencing it at age 5. More striking, I observe that skills do not respond the same way to lack of paternal input.

In terms of nutrition, younger siblings relatively under-perform older siblings in households with an absent father. When both siblings are measured at 5 years old, younger ones are less nourished and are smaller than their older counterparts. Additionally, the relatively lag in height does not vanish over time but on the contrary it persists, when both siblings are measured at 8 and 12 years old. These results hold when pooling the four countries as well as when looking at them separately.[14] Therefore, long-term effects on nutrition suggest that paternal absence in early years leave a stronger mark on the younger sibling relatively to the older one.

According to the World Health Organization, a low height-for-age can be interpreted differently depending on the age of the child, especially in developing countries. Among children aged below 2-3 years old, low height-for-age probably reflects a continuing process of “failing to grow” or “stunting”; for older children, it reflects

[14]Except for Ethiopia at age 12, where the coefficient is marginally positive but statistically not different than zero.

a state of “having failed to grow” or “being stunted” (WHO, 2018).[15] Following this thought, the lower standardized score of the younger sibling with respect to the older one may be explained by a slower growth process when very little. Furthermore, I should expect trends to be parallel if the effect is not cumulative. Indeed, trends in the evolution of anthropometric outcomes are parallel when looking at table D10. My findings provide suggestive evidence of the relevance of critical periods for paternal inputs in nutritional development.

On the contrary, results suggest that cognitive outcomes do not respond to paternal absence the same way. Ethiopian siblings seem to react equally or not to react at all to the absence of the father when the younger experienced it early in life and the older when aged 5. This conclusion holds when measured at 5, 8 and 12 years old. Younger siblings in Peru score relatively higher than their older counterparts in the vocabulary test when 5 years old, but this differential effect vanishes over time. In the case of Vietnam, the estimate is also positive and close to be statistically significant at the 10% level when measured at age 5 (p-value=0.111). These outcomes, especially the ones in Peru and Vietnam, could be explained by either an improvement in cognition by the younger sibling or by a worsening of cognitive skills by the older sibling.

As a reminder, the older sibling is being measured just when the father is gone (i.e. he/she is 5 years old in 2006), hence the relevance of looking at the following rounds for a possible delayed response. Yet, the gap at age 8 and 12 is not statistically different than the one in households with no absent father, both when looking at countries pooled and individually. In terms of magnitudes, the estimates for Peru and Vietnam get closer to zero with respect to the ones when siblings are measured at 5 years old. One interesting case is Vietnam, where the point estimate –while still not statistically significant– actually becomes negative when both siblings are measured at 12. In other words, there could be a reversal in the relatively over-performance of the younger sibling when measured in the short run. However, no strong conclusion can be made from these coefficients as variation is too wide to have precise estimators.

Estimates are in relative terms, thereby my findings could either reflect the responses by the younger sibling or by the older sibling to paternal absence (or a mix of both). For instance, the negative gap in nutrition can be interpreted as a relative improvement of weight and height of the older sibling as a result of paternal absence. Similar reasoning for cognitive outcomes.

Young Lives allows me to perform an ancillary analysis to better understand the main results. Instead of comparing siblings when paternal absence happens during the early years versus when 5 years old, I can compare pairs of siblings who experience absence at an older age, making sure that both siblings had their father present around the age they were born. In brief, the identification strategy resembles the one from the previous exercise except that absence occurs in 2009, when the younger sibling is around 5 years old and the older sibling is 8 years old. Therefore, when they are measured at age 5, only the younger sibling is contemporaneously experiencing the absence at that moment. Details on the identification strategy and econometric specification can be found in section F as well as descriptive statistics (table E11 and E12) and results on estimates (E13, E6, E14 and E15).

When pooling the four countries, I observe that their comparing nutritional and cognitive outcomes do not seem to diverge any differently than the ones of siblings in control households. These results hold when aged 5 (E13) and when siblings grow up (E6). I arrive at a similar conclusion when looking at countries individually (E14), as estimates do not show any (statistically significant) differential response. The only exception is Vietnam, which presents a negative coefficient at age 5 in terms of weight (statistically significant at the 10% level), but not in terms of height. This could be interpreted as a temporary shock in nutrition given that it becomes in-

[15]Global Database on Child Growth and Malnutrition. Retrieved from: www.who.int.

significant the next round and there are no differential responses on height at age 8. If any, the coefficient for height-for-age becomes positive when aged 12. While no clear conclusion can be made from these coefficients, it stresses the consistency of previous results. Specifically, how previous estimates might be capturing something about the first 1,000 days of life, rather than the mere differential effect of sooner versus later absence in nutrition.

This ancillary exercise provides further support on the relevance of the timing of paternal absence, especially in terms of nutritional outcomes. The general results provide strong evidence that paternal absence can be relatively more detrimental to physical development when it occurs during the first years of life and that this negative effect persists over time. One intuitive mechanisms is that paternal absence has a negative income effect on the household and this leads to a reduction in the quantity or quality of food intakes of children. Moreover, this income shock is more harmful when it occurs during the critical periods in the development of children. This hypothesis could be pushed further based on the the fact that common trends between 2002 and 2006 in anthropometrics hold for older sibling in both types of households (see figure 4, section 9 for details). Therefore, it could be argued that the income effect only affects the younger sibling and therefore his/her nutritional outcomes. While this is a strong statement, it can also be backed up by the evidence found when performing the second exercise (absence occurring in 2009), where no impact on nutritional outcomes is observed. If that is the case, I would be capturing the absolute effect of paternal absence. Independently of this latter argument, a negative income shock seems to be the most plausible mechanisms to the relative worsening on nutrition.

The impact of the departure of the father on cognition is not clear when looking at the aggregated results as it appears that its effect is country specific. Cultural norms, as well as heterogeneities in family structures could be behind these outcomes.

In the following section I will explore the two mechanisms mentioned above; the negative income effect which may explain relative deterioration in anthropometrics, and changes in family structure which may give some hints on the unexpected direction of the results on cognition. Moreover, I will explore temporary versus more permanent absence. Due to the limitations on the sample size, most of the analysis on mechanisms will be done pooling the four countries together, with the downside of potentially missing deeper insights.

8 Mechanisms

So far, I study the reduced form of the causal relationship between paternal absence and child development when absence occurs early in life. However, the way fathers affect their offsprings' outcomes is through different inputs they provide in the production function of skills. In this section I attempt to understand what are the factors affected by the absence of the father which seems more plausible to impact on nutrition and cognition. I will focus on the results for the main specification and sample, absence in 2006. Results on the alternative sample can be found in the Appendix.

8.1 Income effect

The hypothesis is that the negative income shock –as a result of the father's departure– could have a relatively bigger repercussion on anthropometrics if it happens when the child is in her first years of life rather than at a later stage. This could be explained by the critical periods in nutritional development during the first 1,000 days of life. The following results provide support for this mechanism.

During the second round (2006), households were asked about their current economic situation and the one four years ago (at baseline, in 2002). Table 4 shows that households that suffer absence in the second round are more likely to report that their economic situation is relatively worse than the one in the previous round. This coefficient is statistically significant at the 1 percent level. On the other hand, these households are less likely to report they are better or that their situation has not changed –while not statistically significantly different than zero. Each regression is controlled by country in order to capture any potential macroeconomic shock.

Table 4: Absence in 2006: Change in economic situation pre and post departure of the father in households with absence in 2006

	(1)	(2)	(3)
VARIABLES	Same	Worse	Better
HH w/absence in 2006	-0.0609	0.0981***	-0.0371
	(0.0377)	(0.0322)	(0.0427)
Observations	2,365	2,365	2,365
R-squared	0.046	0.038	0.026
Observations Control HH	2,212	2,212	2,212
Mean Control HH	.421	.166	.412
Observations Treatment HH	153	153	153
Mean Treatment HH	.359	.261	.379

Notes: Standard errors in parentheses, clustered at the sentinel site level. *** p<0.01, ** p<0.05, * p<0.1. Variables are constructed from the following question in round 2 (2006): "During the period (current/four year ago), how would you describe the household you were living in?" (1) very rich (2) rich (3) comfortable-manage to get by (4) struggle-never have quite enough (5) poor (6) destitute. Variable *same* is equal to 1 if the situation has no changed and 0 otherwise. Variable *worse* is equal to 1 if current situation reported is higher (as the higher the poorer) than the one four years ago and 0 otherwise. Variable *better* is equal to 1 if current situation reported is lower than the one four years ago and 0 otherwise.

In the ancillary exercise from the previous section I show that the gap in development between siblings experiencing absence at age 5 versus 8 was not any different that the gap in development between sibling not experiencing absence at all. One possibility is that, unlike the case in which absence occurs in the early years for the younger sibling, these households do not experience an income shock as a result of paternal departure. This scenario would weaken the strength of my conclusions. In order to provide some insight on this alternative, I perform a similar exercise by looking at expenditures in households experiencing paternal absence in 2009 versus those that did not. Table F16 shows that expenditures evolved less on household experiencing absence with respect to the control households, were the coefficient of interest is statistically significant for food expenditures at the 10% level.[16]

Even if no causality can be claimed, results provide support to the idea that the absence of the father is correlated to a negative income shock. While in both cases- absence in 2006 and in 2009– this income shock was present, the worsening in the economic situation has only a differential impact on nutrition for children in their first years of life when compared to their older siblings.

8.2 Family structure

Family structures may differ across cultures, especially in terms of family size and the presence of extended family members in the household. Results in cognition appear to be counter intuitive, as they do not show the same reaction as nutritional outcomes do to paternal absence. The hypothesis is that family dynamics could compensate the departure of the father, acting as a buffer to the disruptive effect of

[16]Section G.1 shows a detailed explanation of this exercise. As an exploratory exercise, I also compare in table F17 the response of whether the household is economically the same, worse or better in 2006 versus in 2002. In other words, before the departure of the father.

not having a present paternal figure.

I first look at the percentage of families that can be defined as nuclear families[17] in the sample. I find dissimilarities across countries. In Ethiopia, around 70% of the sample can be defined as a nuclear family, in Peru this percentage goes to 62%, in Vietnam 53% while in India it stands at 37%. Table F18 shows the differential in the probability of belonging to a nuclear family between treated households that experienced absence in 2006 (column 1) versus those that did not. In India, Peru and Vietnam, treated households are as likely to belong to a nuclear family as in control households as estimates are not statistically significant.[18] These is not the case for Ethiopia, where the likelihood of belonging to a nuclear family is around 13 percentage points higher than in the control group. Column 2 of table F18 presents results for the study on absence in 2009.

Table 6 presents the difference in the change of the number of adults –excluding parents– from 2002 (baseline) to 2006 between households that experienced the absence in 2006 versus those that did not. I observe that–when pooling the four countries– households with and without an absent father had on average the same number of adults living in the household before paternal departure occurred for the treated homes.[19] Nevertheless, household dynamics appear to differ when the father is not present anymore. As the estimate of the interaction term suggests, in households where the father was absent, the number of other members did not fall as much as in households where the father did not leave (still not statistically significant). As an ancillary exercise, I perform the same analysis for households where absence occurs in 2009 (F19).

Table 5: Absence in 2006: Change in the number of adults over time

	(1)
VARIABLES	Number of adults
HH w/absence in 2006	-0.0163
	(0.141)
year 2006	-0.450***
	(0.0489)
HH w/absence in 2006 * yr 2006	0.215
	(0.139)
Mean HH w/ no absence at year 2002	1.05
Mean HH w/absence at year 2002	.97

Notes: Standard errors in parentheses, clustered at the sentinel site level. *** p<0.01, ** p<0.05, * p<0.1. Adult is defined as other members excluding parents, older than 14 years old by 2002. No household fixed effects were included. Each regression is controlled for country dummies.

A related question is whether time allocation to care-giving among household members changes over time and if so, if this variation is different across household types. Using information on who is the main caregiver of the older child, I look at the whether who the main caregiver is changes over time for both household types. I explore whether this person is the mother, the father or another member in 2002 and 2006. I find that in 2002 the main caregiver tends to be the mother for both type of households. This changes in 2006, but only for treated households, where

[17] I define nuclear family as a household where no other adult (older than 14 years old), except the parents, is reported in the household roster at baseline. Because I include as adult those siblings over 14 years old, this definition differs from the standard one. Nevertheless, I will call this type of households as "nuclear" for simplicity.

[18] Magnitudes can be calculated by adding reference coefficient plus interaction between HH w/absence and each country. Significance was obtain from running individual regressions per country.

[19] Consistent with previous results, the exception being Ethiopia where the number of adults is lower for households with an absent father.

mothers are less likely to be the main caregiver, being replaced by another member in the household.

Table 6: Absence in 2006: Main caregiver over time

	main caregiver		
	(1)	(2)	(3)
VARIABLES	mother	father	other
HH w/absence in 2006	0.00184	-0.000663	-0.00118
	(0.00697)	(0.00690)	(0.000846)
year 2006	-0.00181	0	0.00181
	(0.00265)	(0.00241)	(0.00144)
HH w/absence in 2006 * yr 2006	-0.0570**	-0.00654	0.0636***
	(0.0233)	(0.00693)	(0.0221)
Observations Control	2212	2212	2212
Mean Control in 2002	.99	.01	0
Observations Treatment	153	153	153
Mean Treatment in 2002	.99	.01	0

Notes: Standard errors in parentheses, clustered at the sentinel site level. *** p<0.01, ** p<0.05, * p<0.1. Variable *mother (father)* is equal to 1 if the main caregiver is the mother (father) and 0 otherwise. Variable *other* is equal to 1 if neither the mother nor the father is the main caregiver and 0 otherwise. Main caregiver is defined as follows: (1) in 2002, the three dummies are created based on question "What is your relationship to the child?" when doing the questionnaire to the main caregiver (2) in 2006, the three dummies are created based on question "Is this the primary caregiver?" (when asked about mother and father). No household fixed effects were included. Each regression is controlled for country dummies.

So far, I have been looking at the presence of other members apart from parents. However, the fact that mothers are relatively less likely to be the main caregiver in households with an absent father in 2006 raises the question if mothers are also gone and whether results hold when exploring this. Table F21 shows the total number of cases of maternal absence in 2002 and 2006 in households with and without an absent father in 2006. As the table indicates, maternal absence in 2002 was rare in control households and null in treated households. However, absence in 2006 was relatively more common in treated households and particularly salient in India and Vietnam [20]. Keeping in mind the limitations in terms of sample size, I replicate the main results on table F22 including an interaction term for being the younger sibling in a household with an absent father and absent mother in 2006[21]. I find consistency in the results across rounds.

As a final exercise, I replicate the main results but including an interaction term with a dummy indicating if the household was a nuclear family type at baseline regardless of how dynamics changed with the departure of the father. I would expect that in households that were originally composed of only the mother and the father, effects could be potentially different than when other members were around from the beginning. Table F20 presents estimates when 5, 8 and 12 years old, for the sample which includes paternal absence in 2006. I will focus on cognitive outcomes. Estimates show that in households in which other adults lived with the older sibling at baseline, the relative effect of paternal absence was positive at each age round. However, when interacting with the fact of belonging to a nuclear family at baseline, the coefficient decreases or even becomes negative in the long run. This is consistent with the idea that those households in which initially there were no other adults members had a harder time compensating for the absence of the father, independently if other members arrived when the father was gone.

From these results I find suggestive evidence that, with the departure of the

[20] 12 cases (36%) in India and 4 cases (15%) in Vietnam)

[21] Please note that, unlike variable "HH w/absence in 2006" which represent households with an absent father in 2006 but not in 2002, variable "mother absent in 2006" is a dummy variable for maternal absence in that year regardless of what happened in 2002 with the mother.

father, households in the treatment group changed their family dynamics differently than households where the father was present. Once again, I cannot claim causality between the departure of the father and the differential in the changes in family composition. However, this evidence may provide a plausible explanation on the dissimilar response in terms of nutrition and cognition siblings have when the father is absent. Other members in the households are able to mitigate lack of paternal input. What is more interesting is the findings when absence happens later on in the siblings' life. First, the same correlation between absence and changes in family structure is found among households where absence occurs when siblings are aged 5 and 8. However, this seems not to translate to a gap in cognitive development. While purely speculative, this could suggest higher sensitivity during the early years when it comes to the ability of other caregivers to mitigate risk factors.

8.3 Enrollment to education and time use at 5 years old

One possible explanation for the non-negative estimate (particularly for Peru and Vietnam) on cognition is that the absence of the father increased the incentives to send children to childcare centers. This could be due to a higher demand for external childcare if the caregiver works outside of home. The hypothetical higher probability that the younger sibling started education at an earlier stage than the older sibling could be translated into a relative improvement in cognitive skills.

The household survey includes a question on time use for all children in the household aged 5 to 17.[22] I use a dummy variable which takes value 1 if child spent at least 1 hour at school and 0 otherwise. It is important to note that this question is asked to those children 5 years and older, so the analysis is on a smaller sample when analyzing pairs of siblings at age 5 (as some of them are below that age). I exclude India, given that results of cognition do not include this country.

From column 1 on table F23, I find that younger siblings are more likely to spend at least one hour at school when aged 5 versus older siblings when they were the same age in household with an absent father. When looking at siblings aged 8 and 12, I encounter that spending at least one hour in school is the same among siblings and across household types.

If I take the results at 5 years old, this goes in line with the hypothesis that more education can be explaining the relative improvement in cognition in the short run, but that it disappears in the long run.

A second step is to analyze how children spend their time at home. One hypothesis is that younger siblings spend more time in activities that stimulate their cognitive skills versus their older siblings in households where the father was absent. As in the previous exercise, I use questions on time use for children aged 5-17 in the household, which means that I am looking at a sub-sample when analyzing the first period (when both are "around 5 years old"). Time use questions include: (1) productive use (At school (including traveling time to school) and studying outside of school time (doing homework, extra tuition) (2) play (play time / general leisure(including time taken eating ,drinking and bathing)) (3) work (caring for others (younger siblings, ill household members)), domestic tasks (fetching water, firewood, cleaning, cooking, washing, shopping, etc), tasks on family farm, cattle herding, other family business, shepherding (not just farming)), paid (remunerated) work or activities outside of the household or for someone not in the household) and (4) sleeping.

Table 7 below shows the gap in time spent on different activities between older and younger siblings with and without an absent father in the early years. Having

[22] The question is the following "'Typically how many hours did each child in the household (aged between 5 and 17yrs) spend on the following activities during a typical day (from Monday to Friday) in the last week?: At school (including travelling time to school)."'

in mind the non-negative coefficients on cognition for Peru and Vietnam, I will focus on the results for both countries. For Peru, I find no strong evidence that the result can be explained by a differential investment in productive activities. However, this is the case for Vietnam, where there is suggestive evidence of a relative increase in time spent on productive activities (but not statistically significant at age 5 and a reduction in leisure time (significant at the 5% level).

Table 7: Time use outcomes when 5 years old between younger and older siblings among households with and without and without absence in 2006

VARIABLES	spend time at school	Time use: productive	play	work	sleep
	(1)	(2)	(3)	(4)	(5)
Ethiopia					
Younger sibling	0.0271	0.101	1.039*	-0.691***	0.257
	(0.0327)	(0.179)	(0.523)	(0.234)	(0.194)
HH w/absence in 2006 * younger sibling	0.144	0.576	-2.199**	1.023**	0.0157
	(0.116)	(0.635)	(0.911)	(0.473)	(0.194)
Observations Control	313	313	313	313	313
Mean Control	.13	.9	9.37	1.72	10.76
Observations Treatment	29	29	29	29	29
Mean Treatment	.07	.55	10.48	.72	10.86
India					
Younger sibling	-0.0364	1.432***	-0.504	0.0923	0.0293
	(0.0417)	(0.380)	(0.552)	(0.219)	(0.301)
HH w/absence in 2006 * younger sibling	0.0877**	0.716	0.0496	0.0440	0.616**
	(0.0361)	(0.455)	(0.761)	(0.131)	(0.242)
Observations Control	435	435	435	435	435
Mean Control	.93	6.88	5.71	.45	9.84
Observations Treatment	25	25	25	25	25
Mean Treatment	.92	6.92	5.08	.2	9.32
Peru					
Younger sibling	0.316	1.357	3.651***	-0.822	-1.351
	(0.264)	(2.152)	(0.868)	(0.638)	(0.928)
HH w/absence in 2006 * younger sibling	0.112	0.141	0.214	-0.449	0.494
	(0.101)	(0.638)	(0.629)	(0.563)	(0.304)
Observations Control	383	383	383	383	383
Mean Control	.79	4.5	3.81	1.28	10.1
Observations Treatment	29	29	29	29	29
Mean Treatment	.76	4.48	4.21	1.76	10.1
Vietnam					
Younger sibling	0.0832	0.475	1.940	0.962	-0.707
	(0.359)	(3.293)	(1.663)	(1.828)	(0.456)
HH w/absence in 2006 * younger sibling	0.0957	0.820	-1.638**	0.0699	0.213
	(0.110)	(0.527)	(0.762)	(0.324)	(0.385)
Observations Control	191	191	191	191	191
Mean Control	.91	5.71	7.86	.37	9.98
Observations Treatment	14	14	14	14	14
Mean Treatment	.86	4.14	9.5	.14	10.21

Notes: Standard errors in parentheses, clustered at the sentinel site level. *** p<0.01, ** p<0.05, *p<0.1. Variables were constructed from the following question: "'Typically how many hours did each child in the household (aged between 5 and 17yrs) spend on the following activities during a typical day (from Monday to Friday) in the last week?:"'. Categories were grouped as follows: (1) productive use (at school (including traveling time to school) and studying outside of school time (doing homework, extra tuition) (2) play (play time / general leisure(including time taken eating ,drinking and bathing)) (3) work (caring for others (younger siblings, ill household members)), domestic tasks (fetching water, firewood, cleaning, cooking, washing, shopping, etc), tasks on family farm, cattle herding, other family business, shepherding (not just farming)), paid (remunerated) work or activities outside of the household or for someone not in the household) and (4) sleeping. All regressions include a household fixed effect and are controlled for age, gender and birth order.

8.4 Temporary versus more permanent absence

Permanent absence can have a different impact on the development of children than temporary absence. If the father has left the household temporarily, he might still support his family economically, be in touch with his relatives, and his return could lead to a catch up in development of the children. Contrarily, permanent departure could be the result of abandonment by the father, causing a different development path of his offsprings.

While there is no information on the reason why the father was not present in the round under study, I know whether the father was present in the following rounds. I can differentiate temporary versus more permanent absent by distinguishing cases in which the father came back 4 years after the episode of absence. In table F24 I replicate the main results by interacting the the term of interest with a dummy equal to 1 if the father was still absent in 2009 and 0 otherwise. The upper panel presents results when children are 5 years old while the lower panel presents results pooling all years together.

When looking at height-for-age, which is the most robust and persistent effect, I observe that the gap in nutritional performance is not statistically different when absence is temporary versus when it is more permanent. This is the case in both panels. When it comes to cognitive skills, it seems that longer periods of absence accentuates the relative over performance of the younger sibling versus the older one when they are five. While purely speculative, this finding could be interpreted as a clearer overcompensating response from households facing a more permanent departure of the paternal figure. However, I do not find any evidence of a differential effect when pooling all years together.

9 Robustness check

9.1 Common trends

The main assumption of this paper is that, if the father had not temporarily left the household, the differences in development outcomes between siblings would have been the same when comparing both types of households (those considered treated and those considered in the control group). This assumption also means that there was no shock that impacted the outcomes of the younger and older sibling differently and also led to the departure of the father at the same time.

Ideally, I could provide support for the assumption by looking at the trends in the outcomes before the departure of the father occurs for both sibling, and compare it with those living in control households. Unfortunately, there is no available information on the outcomes of the younger siblings previous to round 3 (2009). However, there is data on the older sibling as well as household characteristics for the first and second round. This allows me to explore common trends. It is important to note I will be comparing outcomes before the departure occurs and outcomes during paternal absence. In other words, trends cannot be considered as pre-treatment.

I perform a difference in difference approach on nutritional outcomes for the older sibling and household characteristics by comparing their evolution in households that experienced the absence of the father versus those that did not. I look at household size, whether the family lives in the rural area and a dummy equal to 1 if the respondent states there has been a shock that has decreased economic welfare of the households. Additionally, I analyze socioeconomic variables such as a wealth index and its components.[23]

The left panel of figure 4 provides support to a parallel evolution in time across older sibling for households that experienced absence in 2006 versus those that did not. Regarding household characteristics, both type of households do not seem to

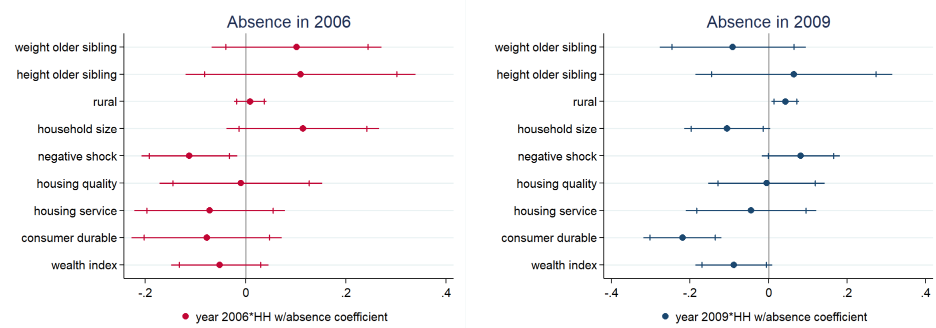
[23]The wealth index is constructed from three equally weighted indices: housing quality, access to services and consumer durables; which I standardize at the country level.

evolve differently in terms of wealth measures. In terms of household size,[24] the positive coefficient is consistent with the previous finding on changes in family structure based on number of adults. Nevertheless, it is not statistically significant. The only dimension in which household types had evolved differently is in terms of the probability of experiencing a negative shock, where the coefficient is negative and statistically significant. The variable includes a variety of shocks such as economic and health shocks (like death or illness), agricultural and natural shocks, birth of a new member[25], etc.

Bottom-line, I find support for the common trend assumption. The only exception is the differential in the probability of experiencing a negative shock. This does not invalidate my conclusions as long as the negative shock is a consequence of the departure of the father. In this way, I would still be capturing the effect of parental absence. A second alternative is that the negative shock is the main source for the results I find and the departure of the father was simply a response to this shock. Hence, the central issue is the establishment of the direction of causality. I tackle this potential threat by testing whether the main conclusions hold once controlling for the probability of having experienced a negative shock in 2006 in section 9.2.

I perform the same exercise for households in which absence occurs in 2009, as it allows me to more accurately study common trends as outcomes are pre-treatment (first and second round). As the panel on the right shows in figure 4, even though children's outcomes evolved in a parallel way, there are reasons to question the common trend assumption. Household experiencing absence seem to do worse in consumer durable index as well as wealth index after the father departure. They also seem to be more likely to have experienced a negative shock and more likely to have moved to a rural area. Moreover, household size has increased less than in the control group.

Figure 4: Common trends: relative change in time across household types



Notes: Standard errors in parentheses, clustered at the sentinel site level. The two panels present the coefficient of interaction term being a household that experienced an absence (in 2006 on the left panel and 2009 on the right panel) and year 2006 (second period). These coefficients are the result of running a regression on the changes across time (2002 vs. 2006) across household types (households with absence and no absence). The figure includes 90% and 95% confidence intervals. Variable *year 2006* is equal to 1 if the outcome is measured in 2006 and 0 if measured in 2002. Variable *year 2006 * HH w/absence* is an interaction term in which HH w/absence is equal to 1 if the household experienced an absence in 2006 (left panel) or in 2009 (right panel) and 0 otherwise.

9.2 Controlling for negative shocks in 2002 and 2006

From the descriptive statistics (table B4) I observe that households experiencing absence are more likely to suffer from a negative shock at baseline (2002) as compared to those that do not experience absence. This could be a threat as long as it reflects selection to paternal departure and therefore different trends. A second finding from the previous sub-section is that the exposure to negative shocks increased less in households experiencing absence. To address these potential threats, I first control

[24]It includes adults and young children in the households –possibly new siblings– and excluding the older sibling under analysis and the father.

[25]I include it given the implications in terms of food expenditures.

for differences in likelihood of experiencing a negative impact in 2002 and then in 2006. With the first exercise, I attempt to address the higher vulnerability at baseline of treated households by controlling the main results for the exposure to shocks in 2002. In this way, I compare families more similar in terms of vulnerability to negative events at baseline. The second exercise will allow me to assess to what extent the changes in exposure to negative shocks across type of households might explain the differences in development between siblings.

Table G25 shows estimates including interactions with a dummy variable indicating if the household reports to have experienced a negative shock in 2002.[26] In this table all years were pooled together (weight-for-age having data for the first two years). The table provides some evidence that most of the impact is explained by most vulnerable households at baseline. In terms of cognition, the coefficient with the interaction might suggest that the positive gap on PPVT test are mainly explained by less vulnerable households. From these findings I find support to my main results as estimates on weight, height and cognition seem to be robust to the fact that treated households were more vulnerable to shocks at baseline.

Table G26 shows that the impact of having an absent father interacted with the occurrence of a negative shock in 2006. I do not find strong arguments to believe that children in households suffering a negative shock in 2006 are more or less exposed to the absence of the father. This is particularly true for height for age and cognition, where the magnitude of the interaction term of interest with having experienced a negative shock in 2006 is close to zero.

9.3 Results with alternative control groups

In all the previous results, the control group had no further restrictions but the fact that the father had to be present in the rounds under analysis. However, no constraints were imposed in terms of his presence in the following rounds. There are reasons to believe that this could not be the most suitable control group. First, I am including households in which no absence occurred afterward, whose characteristics could differ from the treated households in such a way that the common trend assumption is threaten. Hence, it would be more appropriate to compare treated households with families in which the father was absent at some point in the future. On the contrary, having households with episodes of absence afterward could also have its downside. When comparing the relative gap between siblings in treated versus control households when the latter actually experiences the absence in the following rounds, I might be capturing the effect of different timing of paternal departure and its effects on the relative gap of siblings.

As an exploratory exercise, I replicate the main specification with alternative control groups. For each outcome, the first column presents the results previously interpreted as a reference, the second includes a control group in which the father was present in each round while in the third one the control group was constructed with all households in which absence occurred at some point in the future.

Table G27 presents results when siblings are measured at age 5, 8 and 12 years old and all years together, when absence occurs in 2006.[27] The interaction term maintains its negative sign both for weight and height for age regardless of the control group with which households with an absent father in 2006 are compared. It is interesting to note that the effect seems to be stronger when compared the gap at age 5 between treated and control group when the latter includes those households that experienced absence at some point in the following rounds. This type of households

[26] The dependent variable negative shocks includes a long list of different shocks, which are self-reported and some of them may not be particularly relevant for the development of children. However, I decided to add all of them in order not to make any arbitrary decision.

[27] A similar exercise is performed for sample in which absence occurs in 2009. Table G28 shows the results.

could be considered better suited as a comparison group, given the propensity to experience absence of the father at some point. Nevertheless, confidence intervals seem to overlap, providing support to the fact that effects are statistically the same regardless of the control group chosen. Similar conclusions apply for cognitive skills.

10 Conclusion

In this paper I attempt to estimate the causal effect of paternal absence during the first years of life on early childhood development. I take advantage of Young Lives Study, a unique dataset that provides longitudinal data on early childhood development in countries so diverse as Ethiopia, India, Peru and Vietnam. Using skill outcomes of pair of siblings of similar age at different points in time, I take advantage of variations in paternal absence just after born versus at age 5.

I find striking evidence of dissimilar implications of paternal absence across skills when departure occurs at the beginning of a child's life. I show that the absence of the father very early in life has a negative effect on nutritional development both in the short and long run with respect to an absence at a later stage in life. While the negative and significant effect persists on height, it does not increase nor decrease over time. Moreover, these conclusions hold across the four countries. Interestingly, the response on cognition goes in a different direction. Estimates indicate that the effect of early absence has a null average effect when pooling countries together. When desegregating them, I find a positive effect for Peru and Vietnam– being significant for the former one– while statistically null for Ethiopia.

I go one step further and try to explore possible mechanisms to explain my results. Unfortunately, the relatively small sample size limits the extent to which I can make strong conclusions about the findings. Additionally, I face a clear constraint when it comes to performing heterogeneous effects, especially relevant as I am including four different countries with their own contexts. Nevertheless, I find suggestive evidence that households that suffer the departure of the father are more likely to report that their economic situation is relatively worse with respect to baseline versus those households that did not suffer from it. On the other hand, I observe that in households where the father was absent, the number of adult members did not decrease as much as in households where the father did not leave. Additionally, there seems to be a change in resources reallocation, as mothers become less likely to be the main caregiver while other members take their place. While not in all countries, I find in Vietnam some evidence of an increase in productive activities (not statistically significant) and a reduction in leisure (statistically significant) by the younger sibling relatively to the older one in households with an absent father. This could suggest that caregivers invest more time in stimulating skills with the younger sibling when the father is gone.

While these are mere correlations, it provides some intuition on what could be behind the different responses in skill outcomes. I argue that anthropometric outcomes seem to be mainly affected by nutritional inputs and health. These inputs are most likely vulnerable to the income shock associated to the departure of the father. This explains the negative sign on weight and height. On the other hand, cognition may not only be affected by these inputs but also by other factors such as home environment and time investment. If father's departure leads households to reallocate their resources and change their investment decisions, skill responses are expected to be heterogeneous. In addition, certain risk factors– such as lack of time investment by the father– could potentially be compensated by other family members. Indeed, these mechanisms could explain the relative improvement in vocabulary tests. To summarize, while the father appears to be the main financial support of the household, his role in fostering cognitive skills may be substituted by other adults in the household.

An interesting finding comes from looking at heterogeneous effects between households that were more vulnerable at baseline and between households that were more vulnerable to contemporaneous negative shocks the the paternal absence. The analysis reveals that more vulnerable households are more exposed to the negative effect correlated to the departure of the father and are less capable of compensate the potentially disruptive effect of an absent father. However, shock occurring while absence is happening does not seem to make children more or less responsive to paternal absence.

My research is relevant for several reasons. First of all, unlike most of the literature, it focuses on the consequences of paternal rather than maternal inputs on early childhood development outcomes. Second, the use of siblings and household fixed effects overcomes endogeneity issues in this type of studies. Third, it contributes to the literature on the relative importance of the first 1,000 days in human capital accumulation. I analyze nutritional, cognitive and socio-emotional skills, allowing me to study how each ability reacts to the absence of the father. In addition, I measure those skills both in the short and the long run. Moreover, the use of data from developing countries contributes to a broader evidence on the role of the father at an early stage outside the developed world. Lastly, the study of countries so diverse as Ethiopia, India, Peru and Vietnam, provides external validity of my results.

From a policy perspective, this paper provides some useful insight on the importance of taking into consideration the role of the father on early childhood interventions. Ignoring his role in skill formation could potentially limit the impact of these type of policies. Furthermore, understanding the importance of the father in the early years is not only relevant for early childhood interventions but also for policies in the field of labor and gender inequality. Paternal involvement is linked to intra-household resource allocation and the division of responsibilities among the members. The optimality of these decisions is a subject of interest. Moreover, paternal involvement on the child's life has consequences on female labor participation and therefore on gender equality and the mother's well-being. A more balanced share of responsibilities between parents could eventually lead to a reduction in the gender wage gap.

While my paper gets closer to understanding the role of the father in early childhood development, there is still much to learn. Further research must be done on the specificities of the human capital production function. Looking at absence could not be enough as the quality of interactions between the father and the child also matters. Also interesting is to learn what kind of activities fathers perform with their children and if they are any different from the ones performed with the mother. A better understanding of how abilities are formed is the starting point to design efficient policies to reduce skill gap across socio-economic groups and fight against the intergenerational transmission of poverty.

References

Acosta, P. (2006), 'Labor supply, school attendance, and remittances from international migration: The case of El Salvador', *World Bank Policy Research Working Paper* (3903).

Adams Jr, R. H. & Cuecuecha, A. (2010), 'Remittances, household expenditure and investment in Guatemala', *World Development* **38**(11), 1626–1641.

Amuedo-Dorantes, C., Georges, A. & Pozo, S. (2010), 'Migration, remittances, and children's schooling in Haiti', *The Annals of the American Academy of Political and Social Science* **630**(1), 224–244.

Antman, F. M. (2012), 'Gender, educational attainment, and the impact of parental migration on children left behind', *Journal of Population Economics* **25**(4), 1187–1214.

Attanasio, O., Cattan, S., Fitzsimons, E., Meghir, C. & Rubio-Codina, M. (2020), 'Estimating the production function for human capital: results from a randomized controlled trial in Colombia', *American Economic Review* **110**(1), 48–85.

Attanasio, O. P., Fernández, C., Fitzsimons, E. O., Grantham-McGregor, S. M., Meghir, C. & Rubio-Codina, M. (2014), 'Using the infrastructure of a conditional cash transfer program to deliver a scalable integrated early child development program in Colombia: cluster randomized controlled trial', *BMJ* **349**, g5785.

Bandura, A. (1993), 'Perceived self-efficacy in cognitive development and functioning', *Educational Psychologist* **28**(2), 117–148.

Barham, T., Macours, K. & Maluccio, J. A. (2013), 'Boys' cognitive skill formation and physical growth: Long-term experimental evidence on critical ages for early childhood interventions', *American Economic Review* **103**(3), 467–71.

Barker, D. J. (1990), 'The fetal and infant origins of adult disease', *BMJ: British Medical Journal* **301**(6761), 1111.

Becker, G. S. (1991), *A Treatise on the Family*, Harvard University Press.

Behrman, J. R. & Rosenzweig, M. R. (2004), 'Returns to birthweight', *Review of Economics and Statistics* **86**(2), 586–601.

Black, M. M., Walker, S. P., Fernald, L. C., Andersen, C. T., DiGirolamo, A. M., Lu, C., McCoy, D. C., Fink, G., Shawar, Y. R., Shiffman, J. et al. (2017), 'Early childhood development coming of age: science through the life course', *The Lancet* **389**(10064), 77–90.

Black, S. E., Devereux, P. J. & Salvanes, K. G. (2005), 'The more the merrier? the effect of family size and birth order on children's education', *The Quarterly Journal of Economics* **120**(2), 669–700.

Boyden, J. (2011), 'Young lives: an international study of childhood poverty: Round 3, 2009 [data collection] 2nd edition. uk data services. sn:6853, http://doi.org/10.5255/ukda-sn-6853-2'.

Boyden, J. (2014), 'Young lives: an international study of childhood poverty: Round 2, 2006. [data collection]. 2nd edition. uk data service. sn:6852, http://doi.org/10.5255/ ukda-sn-6852-2'.

Boyden, J., Woldehanna, T., Galab, S., Sanchez, A., Penny, M. & Duc, L. (2016), 'Young lives: an international study of childhood poverty: Round 4, 2013-2014. [data collection]. uk data service. sn: 7931, http://doi.org/10.5255/ukda-sn-7931-1'.

Calero, C., Bedi, A. S. & Sparrow, R. (2009), 'Remittances, liquidity constraints and human capital investments in Ecuador', *World Development* **37**(6), 1143–1154.

Carneiro, P. M. & Heckman, J. J. (2003), 'Human capital policy', *NBER Working Paper* (w9495).

Chae, S. (2016), 'Parental divorce and children's schooling in rural Malawi', *Demography* **53**(6), 1743–1770.

Cools, S., Fiva, J. H. & Kirkebøen, L. J. (2015), 'Causal effects of paternity leave on children and parents', *The Scandinavian Journal of Economics* **117**(3), 801–828.

Cox Edwards, A. & Ureta, M. (2003), 'International migration, remittances, and schooling: evidence from El Salvador', *Journal of Development Economics* **2**(72), 429–461.

Cunha, F. & Heckman, J. (2007), 'The technology of skill formation', *American Economic Review* **97**(2), 31–47.

Cunha, F., Heckman, J. J., Lochner, L. & Masterov, D. V. (2006), 'Interpreting the evidence on life cycle skill formation', *Handbook of the Economics of Education* **1**, 697–812.

Currie, J. & Almond, D. (2011), Human capital development before age five, *in* 'Handbook of labor economics', Vol. 4, Elsevier, pp. 1315–1486.

Dearden, K., Crookston, B., Madanat, H., West, J., Penny, M. & Cueto, S. (2013), 'What difference can fathers make? early paternal absence compromises peruvian children's growth', *Maternal & child nutrition* **9**(1), 143–154.

Doyle, O., Harmon, C. P., Heckman, J. J. & Tremblay, R. E. (2009), 'Investing in early human development: timing and economic efficiency', *Economics & Human Biology* **7**(1), 1–6.

Eickmann, S. H., Lima, A. C., Guerra, M. Q., Lima, M. C., Lira, P. I., Huttly, S. R. & Ashworth, A. (2003), 'Improved cognitive and motor development in a community-based intervention of psychosocial stimulation in northeast Brazil', *Developmental Medicine and Child Neurology* **45**(8), 536–541.

Ekberg, J., Eriksson, R. & Friebel, G. (2013), 'Parental leave—a policy evaluation of the Swedish "daddy-month" reform', *Journal of Public Economics* **97**, 131–143.

Fernald, L. C., Kariger, P., Hidrobo, M. & Gertler, P. J. (2012), 'Socioeconomic gradients in child development in very young children: Evidence from India, Indonesia, Peru, and Senegal', *Proceedings of the National Academy of Sciences* **109**(Supplement 2), 17273–17280.

Fernald, L. C., Weber, A., Galasso, E. & Ratsifandrihamanana, L. (2011), 'Socioeconomic gradients and child development in a very low income population: evidence from Madagascar', *Developmental science* **14**(4), 832–847.

Gertler, P., Heckman, J., Pinto, R., Zanolini, A., Vermeersch, C., Walker, S., Chang, S. M. & Grantham-McGregor, S. (2014), 'Labor market returns to an early childhood stimulation intervention in Jamaica', *Science* **344**(6187), 998–1001.

Ghuman, S., Behrman, J. R., Borja, J. B., Gultiano, S. & King, E. M. (2005), 'Family background, service providers, and early childhood development in the Philippines: Proxies and interactions', *Economic Development and Cultural Change* **54**(1), 129–164.

Giannelli, G. C. & Mangiavacchi, L. (2010), 'Children's schooling and parental migration: Empirical evidence on the 'left-behind' generation in Albania', *Labour* **24**(s1), 76–92.

Grantham-McGregor, S., Cheung, Y. B., Cueto, S., Glewwe, P., Richter, L., Strupp, B., Group, I. C. D. S. et al. (2007), 'Developmental potential in the first 5 years for children in developing countries', *The lancet* **369**(9555), 60–70.

Grantham-McGregor, S. M., Powell, C. A., Walker, S. P. & Himes, J. H. (1991), 'Nutritional supplementation, psychosocial stimulation, and mental development of stunted children: the Jamaican study', *The Lancet* **338**(8758), 1–5.

Haas, L. & Hwang, C. P. (2008), 'The impact of taking parental leave on fathers' participation in childcare and relationships with children: Lessons from Sweden', *Community, Work and Family* **11**(1), 85–104.

Hamadani, J. D., Tofail, F., Huda, S. N., Alam, D. S., Ridout, D. A., Attanasio, O. & Grantham-McGregor, S. M. (2014), 'Cognitive deficit and poverty in the first 5 years of childhood in Bangladesh', *Pediatrics* **134**(4), e1001–e1008.

Hoddinott, J., Alderman, H., Behrman, J. R., Haddad, L. & Horton, S. (2013), 'The economic rationale for investing in stunting reduction', *Maternal & child nutrition* **9**, 69–82.

Hu, F. (2012), 'Migration, remittances, and children's high school attendance: The case of rural China', *International Journal of Educational Development* **32**(3), 401–411.

Huttly, S. & Jones, N. (2014), 'Young lives: an international study of childhood poverty: Round 1, 2002.[data collection] 5th edition. uk data service. sn:5307, http://doi.org/ 10.5255/ukda-sn-5307-2'.

Laajaj, R. & Macours, K. (2017), 'Measuring skills in developing countries'.

Lu, Y. & Treiman, D. J. (2011), 'Migration, remittances and educational stratification among blacks in apartheid and post-apartheid South Africa', *Social Forces* **89**(4), 1119–1143.

Macours, K., Premand, P., Schady, N. & Vakis, R. (2015), 'Experimental evidence from an early childhood parenting intervention in Nicaragua', *Society for Research on Educational Effectiveness* .

Macours, K. & Vakis, R. (2010), 'Seasonal migration and early childhood development', *World development* **38**(6), 857–869.

Nepomnyaschy, L. & Waldfogel, J. (2007), 'Paternity leave and fathers' involvement with their young children: Evidence from the American ECLS-B', *Community, Work & Family* .

Nguyen, C. V. (2016), 'Does parental migration really benefit left-behind children? comparative evidence from Ethiopia, India, Peru and Vietnam', *Social Science & Medicine* **153**, 230–239.

Nguyen, C. & Vu, L. (2016), 'Should parents work away from or close to home? the effect of parental absence on children's time use in Vietnam', *Journal of Human Development and Capabilities* **17**(1), 110–124.

Outes-Leon, I., Porter, C. & Sanchez, A. (2011), 'Early nutrition and cognition in Peru', *IDB Working Paper No. IDB-WP-241, Inter-American Development Bank (IDB), Washington, DC* .

Pathak, Y. & Macours, K. (2017), 'Women's political reservation, early childhood development, and learning in India', *Economic Development and Cultural Change* **65**(4), 741–766.

Paxson, C. & Schady, N. (2007), 'Cognitive development among young children in Ecuador the roles of wealth, health, and parenting', *Journal of Human resources* **42**(1), 49–84.

Rege, M. & Solli, I. F. (2013), 'The impact of paternity leave on fathers' future earnings', *Demography* **50**(6), 2255–2277.

Rosenberg, M. (1965), 'Self esteem and the adolescent.(economics and the social sciences: Society and the adolescent self-image)', *Science* **148**, 804.

Rotter, J. B. (1966), 'Generalized expectancies for internal versus external control of reinforcement.', *Psychological monographs: General and applied* **80**(1), 1.

Rubio-Codina, M., Attanasio, O., Meghir, C., Varela, N. & Grantham-McGregor, S. (2015), 'The socioeconomic gradient of child development: Cross-sectional evidence from children 6–42 months in Bogota', *Journal of Human Resources* **50**(2), 464–483.

Sanchez, A., Melendez, G. & Behrman, J. (2016), 'The impact of the juntos conditional cash transfer programme in Peru on nutritional and cognitive outcomes: Does the age of exposure matter?', *Young Lives Working Paper 153* .

Sanchez, A., Penny, M., Woldehanna, T., Galab, S., Boyden, J. & Duc, L. T. (2018), 'Young lives: an international study of childhood poverty: Round 5, 2016. [data collection]. uk data service. sn: 8357, http://doi.org/10.5255/ukda-sn-8357-1'.

Schady, N. (2011), 'Parents' education, mothers' vocabulary, and cognitive development in early childhood: Longitudinal evidence from Ecuador', *American Journal of public health* **101**(12), 2299–2307.

Schady, N., Behrman, J., Araujo, M. C., Azuero, R., Bernal, R., Bravo, D., Lopez-Boo, F., Macours, K., Marshall, D., Paxson, C. et al. (2015), 'Wealth gradients in early childhood cognitive development in five latin american countries', *Journal of Human Resources* **50**(2), 446–463.

Shonkoff, J. P., Garner, A. S., Siegel, B. S., Dobbins, M. I., Earls, M. F., McGuinn, L., Pascoe, J., Wood, D. L., on Psychosocial Aspects of Child, C., Health, F., Committee on Early Childhood, A., Care, D. et al. (2012), 'The lifelong effects of early childhood adversity and toxic stress', *Pediatrics* **129**(1), e232–e246.

Tanaka, S. & Waldfogel, J. (2007), 'Effects of parental leave and work hours on fathers involvement with their babies: Evidence from the millennium cohort study', *Community, Work and Family* **10**(4), 409–426.

Thiombiano, B. G., LeGrand, T. & Kobiané, J.-F. (2013), 'Effects of parental union dissolution on child mortality and child schooling in Burkina Faso', *Demographic Research* **29**, 797–816.

Victora, C. G., De Onis, M., Hallal, P. C., Blössner, M. & Shrimpton, R. (2010), 'Worldwide timing of growth faltering: revisiting implications for interventions', *Pediatrics* **125**(3), e473–e480.

Appendix

A General

A.1 Age of siblings

Age in months was calculated as follows. As a general rule, I chose information from earlier rounds in order to limit recall issues. In the case of Peru, it was calculated based on the birthdate reported in 2009 or 2013. If unavailable, it was calculated based on the age in months reported in the last round (2016). For Ethiopia, ages in months were directly imputed from the dataset as they were included in round 3 and 5. For Vietnam, ages in months were calculated based on birthdates reported in 2009 or ages in years reported the same round or 2013 if unavailable (multiplied by 12). For India, age in months is calculated based on birthdates reported in the first or second round depending on the reliability of dates when compared to age reported in years. If birthdates are not available, then I impute age in years from the third round and for 2016 I impute age in months reported in that round or age in years from the third round (multiplied by 12).

Table A1: Age distribution: older and younger sibling, 2006 and 2009 respectively

	Younger sibling in 2009							
Older sibling in 2006	1	2	3	4	5	6	7	Total
4	0	3	77	189	218	83	0	570
5	2	54	263	429	662	498	5	1,913
6	0	2	8	9	7	9	0	35
Total	2	59	348	627	887	590	5	2,518

Notes: This table reports the age distribution between the older sibling in 2006 and the younger sibling in 2009. Age in months is rounded. Age is calculated from birth dates if available/reliable or age reported in the family roster otherwise.

A.2 Psychosocial skills

Table A2: Psychosocial skills in Young Lives

	Round 3(index children) and round 4 (sibling)	Reversed	Agency Index (all)	Agency Index (excl. work)	Agency Index (excl. school/work)
Agency Index	1. If I try hard, I can improve my situation in life	No	X	X	X
	2. Other people in my family make all the decisions about how I spend my time	Yes	X	X	X
	3. I like to make plans for my future studies and work	No	X	X	X
	4. If I study hard at school, I will be rewarded by a better job in the future	No	X	X	
	5. I have no choice about the work I do – I must do this sort of work	Yes	X		
Pride Index	1. I am proud of my shoes or of having shoes	No	X	X	X
	2. I am proud of my clothes	No	X	X	X
	3. I am never embarrassed because I do not have the right books, pencils or other equipment	No	X	X	
	4. I am proud that I have the correct uniform	No	X	X	
	5. I am proud of the work I have to do	Yes	X		

B Descriptive Statistics

Table B3: Number of households per household type: absence in 2006 and absence in 2009

	Present in 2006	Absent in 2006	Total	Present in 2009	Absent in 2009	Total
Ethiopia	706	56	762	669	37	706
India	636	33	669	568	68	636
Peru	496	37	533	441	55	496
Vietnam	374	27	401	352	22	374
Total	2,212	153	2,365	2,030	182	2,212

Notes: This table reports the number of households in each subsample per country: (1) households that experienced the absence of the father in 2006 (but not in 2002) and those in which the father was present in both rounds (2) households that experienced the absence of the father in 2009 (but not in 2002 nor 2006) and those in which the father was present in the three rounds.

Table B4: Descriptive Statistics at baseline (2002): Households with an absent father in 2006

	Absent father in 2006			No absent father in 2006			Diff	P-value
	N.	Mean	S.D.	N.	Mean	S.D.		
Older sibling's characteristics								
Age (in years)	153	0.96	0.28	2,212	0.97	0.29	-0.01	0.56
Female	153	0.43	0.50	2,212	0.49	0.50	-0.05	0.10
Birth order	153	2.20	1.51	2,212	2.25	1.69	-0.05	0.77
Birth weight (grams)	77	2,961.88	609.35	1,098	3,014.77	573.40	-52.89	0.43
Number of antenatal visits	112	5.09	2.68	1,583	4.79	2.47	0.30	0.33
Creche	153	0.05	0.21	2,212	0.04	0.19	0.01	0.54
Mother's characteristics								
Age of the mother	153	25.66	5.69	2,209	24.88	5.38	0.78	0.04
Mother- Primary or less	148	0.76	0.43	2,168	0.77	0.42	-0.01	0.86
Mother- Secondary education	148	0.17	0.38	2,168	0.19	0.39	-0.02	0.60
Mother- Higher education	148	0.07	0.25	2,168	0.04	0.20	0.03	0.18
Mother regularly takes care	152	0.96	0.20	2,209	0.96	0.20	0.00	0.91
Mother-financial support	153	0.63	0.49	2,199	0.63	0.48	-0.00	0.97
Father's characteristics								
Age of the father	153	31.43	7.18	2,205	30.57	7.43	0.86	0.22
Father- Primary or less	148	0.59	0.49	2,154	0.68	0.47	-0.09	0.06
Father- Secondary education	148	0.34	0.48	2,154	0.26	0.44	0.09	0.04
Father- Higher education	148	0.07	0.25	2,154	0.06	0.25	0.00	0.91
Father regularly takes care	152	0.63	0.49	2,211	0.60	0.49	0.02	0.72
Father-financial support	152	0.84	0.37	2,207	0.88	0.32	-0.05	0.41
Father in agriculture	149	0.62	0.49	2,168	0.68	0.47	-0.06	0.29
Household's characteristics								
Household size	153	5.37	2.25	2,212	5.39	2.20	-0.02	0.90
Rural	153	0.67	0.47	2,212	0.72	0.45	-0.05	0.38
Housing quality index (z)	153	0.22	1.05	2,211	-0.10	1.02	0.32	0.01
Housing service index (z)	153	-0.07	0.98	2,211	-0.16	1.02	0.09	0.41
consumer durable index (z)	153	-0.16	0.90	2,204	-0.23	0.90	0.07	0.39
Wealth index (z)	153	-0.00	0.79	2,202	-0.16	0.81	0.16	0.05
Anyone owns the house	153	0.73	0.45	2,211	0.77	0.42	-0.04	0.24
Experienced negative shock	153	0.62	0.49	2,212	0.50	0.50	0.12	0.00

Notes: This table provides descriptive statistics at baseline (2002) of the characteristics of older siblings, mothers, fathers and of the household, between households that experienced the absence of the father in 2006 but not in 2002 and households that did not experience the absence of the father in neither of those rounds. Children characteristics include: age in months, average of girls, birth order of the sibling, reported birth weight, reported number of antenatal visits during pregnancy, and whether the child had ever attended a creche for at least a whole morning, afternoon or evening almost every week. Mother's characteristics include: age of the mother in years and percentage of mother with primary education or less, secondary education and higher education. Education variables were homogenized for the four countries based on the reported number of years of education. Whether the mother regularly takes care of index children comes from the following question in the household roster: "Responsible for looking after NAME on a regular basis" and whether the mother provides financial support comes from the following question "Helped support NAME financially?". Father's characteristics include the same as the mother's plus whether one of the main economic activities is agriculture. Household's characteristics include: household size, whether they live in the rural area and whether anyone in the household owns the house. Moreover, three wealth indices are included (housing quality, access to services and consumer durables) as well as the overall index; all of them standardized at the country level. The percentage of households that experienced a bad event per household type comes from the following question: "Since you found you were pregnant with NAME have there been any big changes or events that decreased the economic welfare of your household?".

Table B5: Absence in 2006: descriptive statistics of outcomes between siblings when 5 years old

	Older Siblings						Younger Siblings						Siblings within household types			
	Absent father in 2006		No absent father in 2006		Diff older siblings		Absent father in 2006		No absent father in 2006		Diff younger siblings		Diff hh absence		Diff hh no absence	
	N.	Mean/(S.D.)	N.	Mean/(S.D.)	(2-4)	P-value	N.	Mean/(S.D.)	N.	Mean/(S.D.)	(8-10)	P-value	(2-8)	P-value	(4-10)	P-value
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
age	153	5.21	2,212	5.24	-0.03	0.22	150	5.21	2,165	5.04	0.17	0.05	-0.00	0.97	0.20	0.00
		(0.32)		(0.35)				(1.07)		(1.04)						
female	153	0.42	2,212	0.49	-0.06	0.06	150	0.47	2,165	0.51	-0.04	0.35	-0.05	0.33	-0.02	0.17
		(0.50)		(0.50)				(0.50)		(0.50)						
weight	153	16.34	2,208	15.87	0.47	0.05	150	16.18	2,151	15.54	0.64	0.32	0.16	0.78	0.33	0.01
		(2.37)		(2.48)				(7.28)		(4.86)						
weight-for-age	153	-1.09	2,208	-1.34	0.24	0.02	149	-1.41	2,149	-1.38	-0.03	0.80	0.32	0.00	0.05	0.14
		(1.01)		(1.06)				(1.30)		(1.19)						
height	153	104.36	2,208	103.28	1.08	0.02	150	103.23	2,147	102.01	1.23	0.05	1.13	0.13	1.28	0.00
		(5.65)		(5.56)				(7.39)		(9.02)						
height-for-age	152	-1.34	2,207	-1.63	0.28	0.00	149	-1.58	2,127	-1.60	0.03	0.79	0.24	0.07	-0.02	0.57
		(1.02)		(1.06)				(1.18)		(1.33)						
ppvt	115	23.39	1,513	24.44	-1.05	0.49	100	37.07	1,260	35.31	1.76	0.63	-13.68	0.00	-10.87	0.00
		(14.17)		(16.09)				(22.65)		(22.92)						
ppvt zscore	115	-0.96	1,513	-0.92	-0.04	0.62	100	-0.36	1,260	-0.45	0.09	0.39	-0.60	0.00	-0.47	0.00
		(0.60)		(0.64)				(0.71)		(0.76)						

Notes: This table reports outcomes statistics of older and younger siblings when aged 5 in households with and without an absent father in 2006. Outcomes are: age in years, weight, weight-for-age (standardized), height, height-for-age (standardized), PPVT and PPVT z-score (standardized). The first four columns exhibit information for the older sibling in households with an absent father (1-2) and households with no absent father (3-4). Columns 7-10 show information for the younger siblings in households with an absent father (7-8) and without (9-10). For each outcome, the number of observations, mean and standard deviations are provided. Columns 5 and 6 show the difference in mean of older siblings between household types and p-value respectively. Similarly, columns 11 and 12 show the difference for younger siblings. Columns 13 and 14 show the difference in outcomes and p-value for older versus younger siblings in households with an absent father. Columns 15 and 16 compare siblings in households with no absent father.

C All years pooled: identification strategy

The econometric specification pooling all years together is represented by equation (2):

$$
\begin{aligned}
\mathrm{y}_{ih} = \alpha + \gamma \text{younger sibling}_{ih} + \\
\zeta(\text{HH w/absence in } 2006_h \times \text{younger sibling}_{ih}) + \\
\eta(8 \text{ yrs old}_h) + \theta(12 \text{ yrs old}_h) + \mathrm{HH}_h + \phi \mathrm{X}_{ih} + \epsilon_{ih} \quad (2)
\end{aligned}
$$

Where y_{iht} represents the average nutritional (height-for-age) and cognitive outcomes when paired-siblings are 5, 8 and 12 years old (all years together). For weight-for-age, y_{iht} represents the average when paired-siblings are 5 and 8. For socio-emotional outcomes, y_{iht} represents the average when paired-siblings are 8 and 12 years old. $younger\ sibling_{ih}$ is equal to 1 if the outcome belongs to the younger sibling and 0 if the outcome belongs to the older sibling. $HH\ w/absence\ in\ 2006_h$ equals to 1 if in the household the father was absent in 2006 (but not in 2002) and 0 if the father was present in both rounds. 8 $yrs\ old_h$ equals 1 if the outcome belongs to the child when she/he is 8 years old and 0 otherwise. 12 $yrs\ old_h$ equals 1 if the outcome belongs to the child when she/he is 12 years old and 0 otherwise. HH_h is the household fixed effect. X_{ih} includes age, gender and birth order of each child as well as country interactions with each control variable. Standard errors are clustered at the sentinel site level.

D Results

Table C6: Socio-emotional outcomes when 8 and 12 years old between younger and older siblings among households with and without and without absence in 2006

	Self-efficacy (Agency index)			Self-esteem (Pride index)		
VARIABLES	(1)	(2)	(3)	(1)	(2)	(3)
8 years old						
Younger sibling	-0.718*	-0.318	-0.482*	-0.122	-0.441	-0.643***
	(0.402)	(0.219)	(0.244)	(0.293)	(0.282)	(0.243)
HH w/absence in 2006 * younger sibling	-0.00821	0.00695	0.0346	0.00666	-0.0607	0.0388
	(0.176)	(0.142)	(0.140)	(0.224)	(0.131)	(0.108)
Observations Control	954	1709	1924	624	1398	2036
Mean Control	-.04	-.08	-.07	.18	.16	.13
Observations Treatment	66	129	138	39	104	147
Mean Treatment	-.05	-.11	-.09	.28	.15	.05
12 years old						
Younger sibling	0.597***	0.481**	0.443	-0.0897	-0.0343	0.0857
	(0.217)	(0.230)	(0.266)	(0.260)	(0.179)	(0.159)
HH w/absence in 2006 * younger sibling	-0.0853	-0.0538	-0.0290	0.128	0.104	0.0845
	(0.132)	(0.119)	(0.103)	(0.131)	(0.113)	(0.111)
Observations Control	1632	2088	2110	1301	1726	2127
Mean Control	.08	.1	.08	-.08	-.1	-.12
Observations Treatment	120	146	147	102	125	148
Mean Treatment	.08	.11	.1	-.09	-.09	-.14

Notes: Standard errors in parentheses, clustered at the sentinel site level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Each regression includes a household fixed effect and presents, the gap in the corresponding outcome between a child and his/her younger sibling in households that experienced the absence of the father in 2006 versus the gap in household that did not experienced his absence that year. While the upper panel exhibits outcomes when both siblings were 8 years old, the bottom one shows them when they are 12 years old. Dependent variables in column 1-3 are the following subindices: (1) agency index including all questions, (2) agency index excluding those related to work and (3) agency index excluding those related to work and school. Dependent variables in column 4-6 are the following subindices: (1) pride index including all questions, (2) pride index excluding those related to work and (3) pride index excluding those related to work and school. These indices were calculated as explained in section 3. All regressions are controlled for age, gender and birth order. Interactions between each control and each country are included.

Table C7: Nutritional and cognitive outcomes when 5, 8 and 12 years old, coefficients per country

VARIABLES	(1) Weight-for-age	(2) Height-for-age	(3) PPVT Z-Score
5 years old			
Ethiopia			
Younger sibling	0.597***	1.376*	0.0667
	(0.193)	(0.745)	(0.109)
HH w/absence in 2006 * younger sibling	-0.191	-0.0480	-0.0235
	(0.131)	(0.248)	(0.0939)
Number of HH (treatment + control)	762	762	757
India			
Younger sibling	0.382	0.493	
	(0.301)	(0.519)	
HH w/absence in 2006 * younger sibling	-0.439***	-0.190	
	(0.141)	(0.207)	
Number of HH (treatment + control)	669	669	
Peru			
Younger sibling	0.101	1.018***	1.920***
	(0.0866)	(0.170)	(0.184)
HH w/absence in 2006 * younger sibling	-0.0882	-0.426**	0.207*
	(0.189)	(0.158)	(0.103)
Number of HH (treatment + control)	533	533	528
Vietnam			
Younger sibling	-0.705**	-0.366	0.224
	(0.257)	(0.333)	(0.176)
HH w/absence in 2006 * younger sibling	-0.0140	-0.261	0.436
	(0.220)	(0.217)	(0.260)
Number of HH (treatment + control)	401	401	389
8 years old			
Ethiopia			
Younger sibling	0.484	0.863**	-2.260***
	(0.345)	(0.353)	(0.510)
HH w/absence in 2006 * younger sibling	-0.0643	-0.0924	0.0494
	(0.161)	(0.0863)	(0.321)
Number of HH (treatment + control)	761	762	758
India			
Younger sibling	0.363	0.865***	
	(0.262)	(0.265)	
HH w/absence in 2006 * younger sibling	-0.0470	-0.216	
	(0.249)	(0.256)	
Number of HH (treatment + control)	669	669	
Peru			
Younger sibling	0.740	0.200	0.330*
	(0.466)	(0.271)	(0.181)
HH w/absence in 2006 * younger sibling	-0.168	-0.235	-0.0576
	(0.145)	(0.159)	(0.134)
Number of HH (treatment + control)	533	533	532
Vietnam			
Younger sibling	1.185**	1.005**	-2.135***
	(0.440)	(0.374)	(0.416)
HH w/absence in 2006 * younger sibling	-0.253	-0.117	0.0257
	(0.190)	(0.155)	(0.157)
Number of HH (treatment + control)	400	401	398
12 years old			
Ethiopia			
Younger sibling		-0.789	0.452***
		(0.844)	(0.0616)
HH w/absence in 2006 * younger sibling		0.0254	0.0611
		(0.0938)	(0.0418)
Number of HH (treatment + control)		762	663
India			
Younger sibling		0.378**	
		(0.137)	
HH w/absence in 2006 * younger sibling		-0.298	
		(0.222)	
Number of HH (treatment + control)		668	
Peru			
Younger sibling		0.974	-1.180***
		(0.582)	(0.111)
HH w/absence in 2006 * younger sibling		-0.178	0.0360
		(0.190)	(0.0942)
Number of HH (treatment + control)		533	533
Vietnam			
Younger sibling		-0.230	-0.155
		(0.713)	(0.104)
HH w/absence in 2006 * younger sibling		-0.445*	-0.152
		(0.244)	(0.0892)
Number of HH (treatment + control)		401	400

Notes: Standard errors in parentheses, clustered at the sentinel site level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Estimates are per country and outcomes are measured when aged 5, 8 and 12 years old. Each regression includes a household fixed effect. Variable younger sibling is equal to 1 if the individual is the *younger sibling* and 0 if the individual is the relatively older child. Variable *hh with absence in 2006* is equal to 1 if the pair of siblings under comparison belong to a household which experienced the absence in 2006 but not in 2002, and 0 if the father was present in both rounds. Dependent variables in column 1 and 2 are weight-for-age z-score and height-for-age z-score. Both anthropometric z-scores are based on the 2006 WHO child growth standards for children below 5 and 2007 WHO child growth standards for children aged 5-19 years (except for weight-for-age for children older than 10). Dependent variable in column 3 correspond to the standardized score of the Peabody Picture Vocabulary Test (PPVT) at the country level. For this outcome, India is excluded due to lack of data for younger siblings. All regressions are controlled for age, gender and birth order.

Table C8: Socio-emotional outcomes by country when 8 and 12 years old between younger and older siblings among households with and without and without absence in 2006

	Self-efficacy (Agency index)			Self-esteem (Pride index)		
VARIABLES	(1)	(2)	(3)	(1)	(2)	(3)
8 years old						
HH w/absence in 2006 * younger sibling (Ethiopia)	0.334**	0.340***	0.363***	-0.0969	-0.348	0.0374
	(0.167)	(0.106)	(0.0783)	(0.563)	(0.257)	(0.166)
p-value individual regression	[0.064]	[0.005]	[0.000]	[0.867]	[0.199]	[0.827]
HH w/absence in 2006 * younger sibling * India	-1.392**	-0.772**	-0.826**	-0.557	-0.0715	-0.357
	(0.626)	(0.316)	(0.332)	(0.602)	(0.353)	(0.306)
p-value individual regression	[0.101]	[0.171]	[0.176]	[0.007]	[0.105]	[0.238]
HH w/absence in 2006 * younger sibling * Peru	-0.363	-0.198	-0.251	0.492	0.542	0.217
	(0.341)	(0.301)	(0.331)	(0.667)	(0.332)	(0.257)
p-value individual regression	[0.924]	[0.628]	[0.736]	[0.289]	[0.379]	[0.220]
HH w/absence in 2006 * younger sibling * Vietnam	-0.878	-0.660	-0.592	0.0887	0.526	0.125
	(0.565)	(0.424)	(0.390)	(0.690)	(0.328)	(0.281)
p-value individual regression	[0.336]	[0.456]	[0.564]	[0.984]	[0.404]	[0.494]
HH control	954	1709	1924	624	1,398	2,036
HH treatment	66	129	138	39	104	147
P-value younger sibling	.045	.039	.014	.198	.146	.000
P-value HH w/absence in 2006 * younger sibling	.074	.056	.044	.075	.109	.345
12 years old						
HH w/absence in 2006 * younger sibling (Ethiopia)	-0.272	-0.225	-0.193	0.0868	0.00764	0.00333
	(0.226)	(0.223)	(0.191)	(0.230)	(0.225)	(0.190)
p-value individual regression	[0.251]	[0.334]	[0.333]	[0.715]	[0.974]	[0.986]
HH w/absence in 2006 * younger sibling * India	0.772**	0.567**	0.539**	0.00709	0.0522	-0.0532
	(0.307)	(0.281)	(0.238)	(0.420)	(0.357)	(0.340)
p-value individual regression	[0.029]	[0.064]	[0.027]	[0.796]	[0.834]	[0.864]
HH w/absence in 2006 * younger sibling * Peru	0.196	0.142	0.0728	0.0502	0.139	0.200
	(0.327)	(0.311)	(0.280)	(0.342)	(0.302)	(0.292)
p-value individual regression	[0.756]	[0.711]	[0.574]	[0.601]	[0.486]	[0.377]
HH w/absence in 2006 * younger sibling * Vietnam	0.0393	0.0631	0.162	0.119	0.218	0.260
	(0.378)	(0.359)	(0.296)	(0.252)	(0.257)	(0.249)
p-value individual regression	[0.462]	[0.579]	[0.896]	[0.064]	[0.091]	[0.127]
HH control	1,632	2,088	2,110	1,301	1,726	2,127
HH treatment	120	146	147	102	125	148
P-value Sibling	.166	.193	.027	.000	.000	.001
P-value Household type * Sibling	.057	.156	.09	.957	.831	.653

Notes: Standard errors in parentheses, clustered at the sentinel site level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Each regression includes a household fixed effect and presents, by country, the gap in the corresponding outcome between a child and his/her younger sibling in households that experienced the absence of the father in 2006 versus the gap in household that did not experienced his absence that year. While the upper panel exhibits outcomes when both siblings were 8 years old, the bottom one shows them when they are 12 years old. The reference country corresponds to Ethiopia. Dependent variables in column 1-3 are the following subindices: (1) agency index including all questions, (2) agency index excluding those related to work and (3) agency index excluding those related to work and school. Dependent variables in column 4-6 are the following subindices: (1) pride index including all questions, (2) pride index excluding those related to work and (3) pride index excluding those related to work and school. These indices were calculated as explained in section 3. All regressions are controlled for age, gender and birth order. Interactions between each control and each country are included. Each estimate includes in brackets the p-values of the individual regression by country.

Table C9: PPVT score: Sensitivity analysis

VARIABLES	(1) 5 years old	(2) 8 years old	(3) 12 years old
Raw PPVT score			
Younger sibling	12.65**	-42.12***	-6.140
	(5.411)	(9.955)	(4.001)
HH w/absence in 2006 * younger sibling	3.322	0.308	0.154
	(2.392)	(3.810)	(1.182)
Obs. Control	1219	1347	1415
Mean Control (older sibling, no absence)	23.76	68.16	56.59
Obs. Treatment	95	109	110
Mean Treatment (older sibling, absence)	21.46	69.96	56.81
PPVT zscore (conditional mean)			
Younger sibling	0.544*	-0.392	-0.375
	(0.291)	(0.236)	(0.326)
HH w/absence in 2006 * younger sibling	0.112	0.0543	0.0204
	(0.163)	(0.128)	(0.106)
Obs. Control HH	1219	1251	1410
Mean Control (older sibling, no absence)	-.18	.14	.04
Obs. Treatment H	95	102	108
Mean Treatment (older sibling, absence)	-.29	.27	.09
Taking outliers out (values greater than 4sd)			
Younger sibling	0.479**	-1.486***	-0.228
	(0.202)	(0.298)	(0.153)
HH w/absence in 2006 * younger sibling	0.130	-0.0295	0.00669
	(0.0940)	(0.118)	(0.0457)
Obs. Control	1218	1282	1415
Mean Control (older sibling, no absence)	-.89	.57	.28
Obs. Treatment	95	105	110
Mean Treatment (older sibling, absence)	-.97	.72	.32
Same sample across rounds			
Younger sibling	0.540**	-1.952***	-0.295*
	(0.220)	(0.386)	(0.174)
HH w/absence in 2006 * younger sibling	0.0943	0.0710	0.0303
	(0.0969)	(0.191)	(0.0524)
Obs. Control	1106	1106	1106
Mean Control (older sibling, no absence)	-.89	.87	.24
Obs. Treatment	89	89	89
Mean Treatment (older sibling, absence)	-.97	.9	.26

Notes: Standard errors in parentheses, clustered at the sentinel site level. *** p<0.01, ** p<0.05, * p<0.1. The upper panel replicates the main results using the raw score. The upper middle panel replicates the main result using a PPVT zscore in which the standardization is done using a mean that is conditional on age polynomial. The standardization is done for each age group (5, 8 and 12 years old) and per country. The lower middle panel replicates results once outliers (>4sd) are taken out. The lower panel presents results restricting the same sample of households for the three rounds under analysis. All regressions are controlled for age, gender and birth order. Interactions between each control and each country are included.

E Difference in evolution: identification strategy and results

The econometric specification for the difference in evolution is represented by equation (3):

$$\begin{aligned} \text{y}_{iht} &= \alpha + \gamma \text{younger sibling}_{ih} + \delta \text{age 2nd period}_{iht} + \\ &\zeta(\text{HH w/absence in 2006}_h \times \text{younger sibling}_{ih}) + \\ &\eta(\text{HH w/absence in 2006}_h \times \text{age 2nd period}_{iht}) + \theta(\text{younger sibling}_{ih} \times \text{age 2nd period}_{iht}) + \\ &\tau(\text{HH w/absence in 2006}_h \times \text{younger sibling}_{ih} \times \text{age 2nd period}_{iht}) + \text{HH}_h + \phi \text{X}_{iht} + \epsilon_{iht} \end{aligned} \tag{3}$$

Where y_{iht} represents nutritional and cognitive outcomes when paired-siblings are 5 or 8 years old for the first trend and 8 or 12 for the second trend analysis. *younger sibling*$_{ih}$ is equal to 1 if the outcome belongs to the younger sibling and 0 if the outcome belongs to the older sibling. $HH\ w/absence\ in\ 2006_h$ equals to 1 if in the household the father was absent in 2006 (but not in 2002) and 0 if the father was present in both rounds. For the first trend analysis, $age\ 2nd\ period_{iht}$ equals 1 if the outcome belongs to the child when she/he is 8 years old and 0 if the child is 5 years old (first period). For the second trend analysis, $age\ 2nd\ period_{iht}$ equals 1 if the outcome belongs to the child when she/he is 12 years old and 0 if the child is 8 years old (first period). HH_h is the household fixed effect. X_{ih} includes age, gender and birth order of each child. In addition, two variables for age are included (one indicating the age in the first period and another in the second period). The reason for doing so is to avoid putting any restriction on the functional form of variable age. nteractions between each control and each country are included. Again, standard errors are clustered at the sentinel site level. These specifications enable me to assess if the potential effects of having an absent father rather before than after increase or decrease over time.

Table D10: Evolution of nutritional outcomes when 5 and 8 and when 8 and 12 years old between younger and older sibling among households with and without absence in 2006

	5 (1st period) vs. 8 yrs old (2nd period)			8 (1st period) vs. 12 yrs old (2nd period)	
	(1)	(2)	(3)	(1)	(2)
VARIABLES	Weight	Weight-for-age	Height-for-age	Weight	Height-for-age
Younger sibling	0.989**	0.150	0.428*	2.921***	0.649***
	(0.418)	(0.176)	(0.255)	(0.626)	(0.170)
2nd period	-2.634***	0.197	0.996***	-2.982	1.808***
	(0.773)	(0.210)	(0.160)	(1.990)	(0.240)
HH w/absence in 2006 * younger sibling	0.0144	-0.105	-0.205*	-0.438	-0.160**
	(0.553)	(0.0890)	(0.119)	(0.336)	(0.0789)
HH w/absence in 2006 * 2nd period	0.0742	0.0219	-0.106*	0.623	0.0761
	(0.175)	(0.0588)	(0.0613)	(0.410)	(0.0504)
Younger sibling * 2nd period	1.145***	0.365***	0.355***	0.778***	-0.245***
	(0.176)	(0.0545)	(0.0625)	(0.287)	(0.0468)
HH w/absence in 2006 * younger sibling * 2nd period	-0.472	-0.00927	0.0139	-0.966*	-0.0117
	(0.665)	(0.0808)	(0.115)	(0.535)	(0.0847)
Observations	9,108	7,176	8,964	9,095	8,944
R-squared	0.650	0.072	0.080	0.571	0.087
Number of HH	2,277	1,794	2,241	2,275	2,236
Observations Control	2128	1688	2094	2127	2089
Observations Treatment	149	106	147	148	147

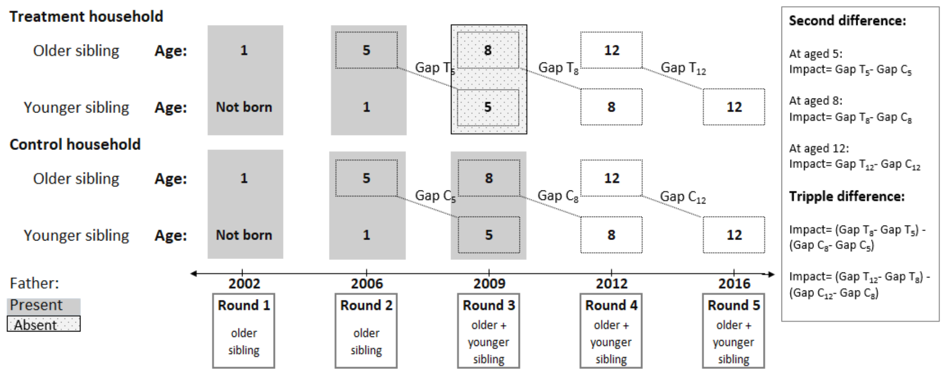
Notes: Standard errors in parentheses, clustered at the sentinel site level. *** p<0.01, ** p<0.05, * p<0.1. Each regression includes a household fixed effect and presents the evolution in the gap in the corresponding outcome between a child and his/her younger sibling in households that experienced the absence of the father in 2006 versus the gap in household that did not experienced his absence that year. The table exhibits in column 1-3 the evolution in nutritional outcomes between two periods, when both siblings are 5 and 8 years old (first and second period respectively). In column 5-6, it shows the evolution when both siblings are 8 and 12 years old (first and second period respectively). Variable *younger sibling* is equal to 1 if the individual is the younger sibling and 0 if the individual is the relatively older child. In column 1-3, variable *2nd period* is equal to 1 when both siblings are 8 and 0 when both siblings are 5 years old. In column 5-6, variable *2nd period* is equal to 1 when both siblings are 12 and 0 when both siblings are 8 years old. Variable *hh with absence in 2006* is equal to 1 if the pair of siblings under comparison belong to a household which experienced the absence in 2006 but not in 2002, and 0 if the father was present in both rounds. The coefficient of interest is variable *hh with absence in 2006 * younger sibling * 2nd period*, which is equal to 1 when the outcome corresponds to the younger sibling in the second period belonging to a household with absence in 2006. Dependent variables are the following: weight, weight-for-age z-score and height-for-age z-score. All anthropometric z-scores are based on the 2006 WHO child growth standards for children below 5 and 2007 WHO child growth standards for children aged 5-19 years (except for weight-for-age for children older than 10). All regressions are controlled for age (first and second period), gender and birth order. Interactions between each control and each country are included.

F Absence in 2009: Paternal absence at age 5 versus at age 8

F.1 Identification strategy

The identification strategy resembles the one from the previous exercise except that absence occurs in 2009, when the younger sibling is around 5 years old and the older sibling is 8 years old. When both siblings are measured at age 5, only the younger one is experiencing paternal absence while the older has not experienced it yet (he will in the following round). Assuming the older sibling has not be treated yet, the hypothesis of this exercise is that having an absent father when around 5 years of life has an effect versus having a present father. When comparing outcomes between siblings when 8 and 12 years old, I test the hypothesis that having an absent father when around 5 years old is different than having it afterward in terms of children's development. Lastly, I test if effects accelerate or not over time.

Figure E5: Absence in 2009: timing of absence and measurement of children's outcomes



Notes: This figure shows the identification strategy for the second exercise. There are two types of households: (1) treated households, the father was present in the first and second round but was absent in 2009 (2) control households, the father was present in the first three rounds. In each household there is information on nutritional, cognitive and socio-emotional outcomes for a pair of siblings, a younger and an older sibling. Therefore, the sample includes those households with at least two siblings. Information on outcomes is only available for the younger sibling from 2009 onward. Two alternative analyses are performed. First, the difference (or gap) between the younger and older sibling in the treatment group when 5, 8 and 12 years old (Gap T_4, T_8 and T_{12} respectively) minus the gap between siblings in the control group (Gap C_4, C_8 and C_{12} respectively). Secondly, the difference in the evolution of the gap between siblings in the treatment group versus the control group when they are 8 versus 5 ((Gap T_8- Gap T_4)-(Gap T_{12}- Gap T_8)) and 12 versus 8 ((Gap T_8- Gap T_4)-(Gap T_{12}- Gap T_8)).

For this exercise I take those paired-siblings in households who had experienced the absence of the father in 2009 (but not before) and compare them against those in households whose fathers were present. As seen in figure E5, the treatment group includes those experiencing the absence in 2009 but not in 2006 nor 2002. The treatment is composed of 182 households while the control comprises 2,030 households (see table B3 for details on sample size per country). Again, the absence affects the younger sibling but unlike the previous exercise, it happens when she/he is around 5 years old, while it affects the older sibling when he/she is around 8 years old. While there is no information on the exact departure date, I know the father was present when the younger sibling was around 1 year old (in 2002).

F.2 Econometric specification

$$y_{ih} = \alpha + \gamma \text{younger sibling}_{ih} + \delta(\text{HH w/absence in 2009}_h \times \text{younger sibling}_{ih}) + \text{HH}_h + \phi X_{ih} + \epsilon_{ih} \quad (4)$$

Where y_{ih} are nutritional (weight-for-age and height-for-age) and cognitive outcomes (standardized PPVT) when paired-siblings are around 5, 8 and 12 years old, while it represents socio-emotional outcomes when children are 8 and 12 years old. $youngersibling_{ih}$ is equal to 1 if the outcome belongs to the younger sibling and 0 if the outcome belongs to the older sibling. $HHw/absence\ in\ 2009_h$ is equal to 1 if the household experiences absence in 2009 (but not in 2006 nor 2002) and 0 if there was no episode of absence in the first three rounds. The coefficient of interest is δ. HH_h is the household fixed effect. X_{ih} includes age, gender and birth order of each child. nteractions between each control and each country are included. 45 Standard errors are clustered at the sentinel site level. With this specification, I examine the contemporaneous impact of having an absent father when around 5 years old versus having a present father until that age. I then explore the medium and long term effect when aged 8 and 12, when both siblings have experienced the absence of the father in the treated households.

F.3 Descriptive statistics

The sample of analysis consists of a total of 2,212 households, in which 182 households experienced the absence of the father in 2009 but not in 2002 nor 2006 (treated households) and 2,030 households that had not experienced the absence of the father neither in 2002, 2006 nor 2009 (households in control group) (see table B3 for details on sample size per country). Table E11 reports information on the characteristics of older siblings, mothers, fathers and their households' characteristics at baseline (i.e 2002). Table E12 shows descriptive statistics of the outcomes of siblings when 5 years old in households with and without an absent father in 2009.

Table E11: Descriptive Statistics at baseline (2002): Households with an absent father in 2009

	Absent father in 2009			No absent father in 2009			Diff	P-value
	N.	Mean	S.D.	N.	Mean	S.D.		
Older sibling's characteristics								
Age (in years)	182	1.00	0.29	2,030	0.97	0.29	0.04	0.10
Female	182	0.53	0.50	2,030	0.48	0.50	0.05	0.16
Birth order	182	2.01	1.44	2,030	2.27	1.71	-0.26	0.08
Birth weight (grams)	95	3,177.16	624.97	1,003	2,999.39	566.19	177.76	0.01
Number of antenatal visits	136	5.24	2.47	1,447	4.75	2.47	0.49	0.05
Creche	182	0.04	0.19	2,030	0.04	0.19	0.00	0.88
Mother's characteristics								
Age of the mother	181	24.12	5.03	2,028	24.95	5.40	-0.83	0.08
Mother- Primary or less	179	0.69	0.46	1,989	0.78	0.42	-0.09	0.05
Mother- Secondary education	179	0.25	0.44	1,989	0.18	0.39	0.07	0.07
Mother- Higher education	179	0.06	0.24	1,989	0.04	0.20	0.02	0.29
Mother regularly takes care	182	0.97	0.16	2,027	0.96	0.20	0.01	0.08
Mother-financial support	182	0.65	0.48	2,017	0.63	0.48	0.03	0.52
Father's characteristics								
Age of the father	181	29.29	7.59	2,024	30.68	7.40	-1.39	0.06
Father- Primary or less	179	0.64	0.48	1,975	0.68	0.47	-0.04	0.33
Father- Secondary education	179	0.27	0.45	1,975	0.26	0.44	0.02	0.66
Father- Higher education	179	0.09	0.29	1,975	0.06	0.24	0.03	0.24
Father regularly takes care	182	0.55	0.50	2,029	0.61	0.49	-0.05	0.25
Father-financial support	182	0.90	0.31	2,025	0.88	0.32	0.01	0.63
Father in agriculture	179	0.69	0.46	1,989	0.68	0.47	0.00	0.91
Household's characteristics								
Household size	182	5.21	2.16	2,030	5.41	2.20	-0.19	0.22
Rural	182	0.66	0.47	2,030	0.72	0.45	-0.06	0.30
Housing quality index (z)	182	-0.10	1.07	2,029	-0.10	1.01	0.00	0.97
Housing service index (z)	182	-0.27	1.05	2,029	-0.15	1.01	-0.12	0.24
consumer durable index (z)	182	-0.20	0.90	2,022	-0.24	0.90	0.04	0.65
Wealth index (z)	182	-0.19	0.83	2,020	-0.16	0.81	-0.03	0.75
Anyone owns the house	182	0.75	0.43	2,029	0.77	0.42	-0.02	0.71
Experienced bad event	182	0.41	0.49	2,030	0.51	0.50	-0.09	0.04

Notes: This table provides descriptive statistics at baseline (2002) of the characteristics of older siblings, mothers, fathers and of the household, between households that experienced the absence of the father in 2009 but not in 2006 nor 2002 and households that did not experience the absence of the father in neither of those years. Children characteristics include: age in years, average of girls, birth order of the sibling, reported birth weight, reported number of antenatal visits during pregnancy, and whether the child had ever attended a creche for at least a whole morning, afternoon or evening almost every week. Mother's characteristics include: age of the mother in years and percentage of mother with primary education or less, secondary education and higher education. Education variables were homogenized for the four countries based on the reported number of years of education. Whether the mother regularly takes care of index children comes from the following question in the household roster: "Responsible for looking after NAME on a regular basis" and whether the mother provides financial support comes from the following question "Helped support NAME financially?". Father's characteristics include the same as the mother's plus whether one of the main economic activities is agriculture. Household's characteristics include: household size, whether they live in the rural area and whether anyone in the household owns the house. Moreover, three wealth indices are included (housing quality, access to services and consumer durables) as well as the overall index; all of them standardized at the country level. The percentage of households that experienced a bad event per household type comes from the following question: "Since you found you were pregnant with NAME have there been any big changes or events that decreased the economic welfare of your household?".

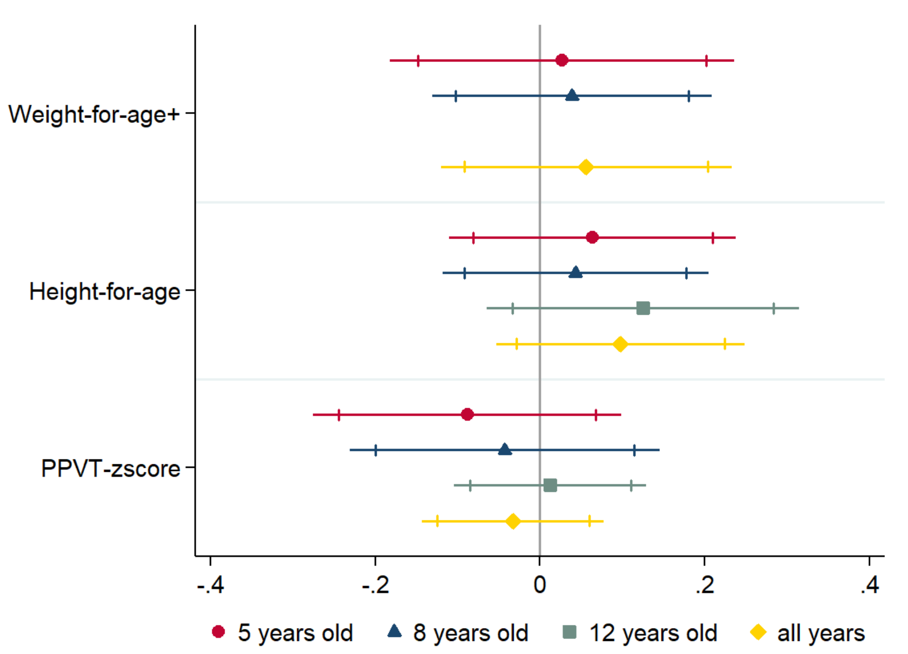
Table E12: Absence in 2009: descriptive statistics of outcomes between siblings when 5 years old

	Older Siblings						Younger Siblings						Siblings within household types			
	Absent father in 2006		No absent father in 2006		Diff older siblings		Absent father in 2006		No absent father in 2006		Diff younger siblings		Diff hh absence		Diff hh no absence	
	N.	Mean/(S.D.)	N.	Mean/(S.D.)	(2-4)	P-value	N.	Mean/(S.D.)	N.	Mean/(S.D.)	(8-10)	P-value	(2-8)	P-value	(4-10)	P-value
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
age	182	5.33	2,030	5.24	0.09	0.00	179	5.01	1,986	5.05	-0.03	0.69	0.32	0.00	0.19	0.00
		(0.35)		(0.35)				(1.09)		(1.04)						
female	182	0.53	2,030	0.48	0.05	0.18	179	0.60	1,986	0.50	0.09	0.03	-0.06	0.22	-0.02	0.28
		(0.50)		(0.50)				(0.49)		(0.50)						
weight	179	16.11	2,029	15.85	0.26	0.36	176	15.52	1,975	15.54	-0.02	0.95	0.59	0.07	0.31	0.01
		(2.79)		(2.45)				(3.02)		(4.99)						
weight-for-age	179	-1.31	2,029	-1.34	0.03	0.78	176	-1.28	1,973	-1.39	0.11	0.21	-0.03	0.76	0.05	0.11
		(1.15)		(1.05)				(1.25)		(1.18)						
height	179	104.01	2,029	103.22	0.79	0.10	175	101.73	1,972	102.03	-0.30	0.70	2.28	0.01	1.19	0.00
		(5.64)		(5.55)				(9.14)		(9.01)						
height-for-age	179	-1.56	2,028	-1.63	0.07	0.42	173	-1.48	1,954	-1.61	0.13	0.21	-0.08	0.39	-0.02	0.67
		(1.04)		(1.06)				(1.36)		(1.33)						
ppvt	108	27.39	1,405	24.21	3.18	0.11	92	39.01	1,168	35.02	3.99	0.22	-11.62	0.00	-10.81	0.00
		(16.60)		(16.04)				(23.70)		(22.84)						
ppvt zscore	108	-0.93	1,405	-0.92	-0.01	0.85	92	-0.44	1,168	-0.45	0.01	0.90	-0.49	0.00	-0.47	0.00
		(0.60)		(0.65)				(0.69)		(0.76)						

Notes: This table reports outcomes statistics of older and younger siblings when aged 5 in households with and without an absent father in 2009. Outcomes are: age in years, weight, weight-for-age (standardized), height, height-for-age (standardized), PPVT and PPVT z-score (standardized). The first four columns exhibit information for the older sibling in households with an absent father (1-2) and households with no absent father (3-4). Column 7-10 show information for the younger siblings in households with an absent father (7-8) and without (9-10). For each outcome, the number of observations, mean and standard deviations are provided. Column 5 and 6 show the difference in mean of older siblings between household types and p-value respectively. Similarly, column 11 and 12 show the difference for younger siblings. Column 13 and 14 show the difference in outcomes and p-value for older versus younger siblings in households with an absent father. Column 15 and 16 compare siblings in households with no absent father.

F.4 Results

Figure E6: Estimates when siblings are 5, 8 and 12 years old for nutritional and cognitive outcomes



Notes: [+]All years together excludes period when children are 12 years old.
This figure reports the estimates of interest (hh with an absent father in 2009 and outcomes for the younger sibling) obtained from the estimation of equation (4), when siblings are 5, 8 and 12 years old and all years together, polling all countries. The figure includes 90% and 95% confidence intervals. Estimates represent the gap in the corresponding outcome between a child and his/her younger sibling in households that experienced the absence of the father in 2009 versus the gap in household that did not experienced his absence that year. Dependent variables are the following: weight (divided by 10 for exhibition purpose), weight-for-age z-score, height-for-age z-score and Peabody Picture Vocabulary Test (PPVT) standardized at the country level. All anthropometric z-scores are based on the 2006 WHO child growth standards for children below 5 and 2007 WHO child growth standards for children aged 5-19 years (except for weight-for-age for children older than 10). For the cognitive outcome, India is excluded due to lack of data for younger siblings. All regressions include a household fixed effect and are controlled for age, gender and birth order. Interactions between each control and each country are included.

Table E13: Nutritional and cognitive outcomes when 5 years old between younger and older siblings among households with and without and without absence in 2009

VARIABLES	(1) Weight-for-age	(2) Height-for-age	(3) PPVT Z-Score
Polling all countries			
Younger sibling	0.161	0.388	0.439**
	(0.259)	(0.347)	(0.204)
HH w/absence in 2009 * younger sibling	0.0265	0.0639	-0.0898
	(0.105)	(0.0873)	(0.0572)
With country interactions			
Younger sibling (Ethiopia)	0.571***	1.358*	0.0898
	(0.197)	(0.745)	(0.106)
	[0.010]	[0.090]	[0.413]
Younger sibling * India	-0.198	-0.957	
	(0.367)	(0.921)	
	[0.251]	[0.475]	
Younger sibling * Peru	-0.435*	-0.334	1.781***
	(0.222)	(0.767)	(0.276)
	[0.210]	[0.000]	[0.000]
Younger sibling * Vietnam	-1.115***	-1.628**	0.111
	(0.338)	(0.808)	(0.207)
	[0.069]	[0.411]	[0.285]
HH w/absence in 2009 * younger sibling (Ethiopia)	0.0969	0.0951	-0.0881
	(0.203)	(0.190)	(0.0934)
	[0.645]	[0.628]	[0.364]
HH w/absence in 2009 * younger sibling * India	0.0941	-0.0650	
	(0.270)	(0.218)	
	[0.303]	[0.787]	
HH w/absnece in 2009 * younger sibling * Peru	-0.165	-0.106	-0.0348
	(0.279)	(0.267)	(0.123)
	[0.731]	[0.956]	[0.145]
HH w/absence in 2009 * younger sibling * Vietnam	-0.434	0.246	0.0597
	(0.272)	(0.317)	(0.185)
	[0.083]	[0.206]	[0.864]
P-value younger sibling	.013	.005	.000
P-value HH w/absence in 2009 * younger sibling	.188	.69	.862
Observations	4,290	4,244	2,438
R-squared	0.077	0.036	0.535
Number of HH	2,145	2,122	1,219
Observations control	1,972	1,952	1,131
Mean control	-1.34	-1.64	-.88
Observations treatment	173	170	88
Mean treatment	-1.31	-1.59	-.94
Includes India	Yes	Yes	No
Age polynomial	No	No	No

Notes: Standard errors in parentheses, clustered at the sentinel site level. *** p<0.01, ** p<0.05, * p<0.1. Each regression includes a household fixed effect and presents the gap in the corresponding outcome between a child and his/her younger sibling in households that experienced the absence of the father in 2009 versus the gap in household that did not experienced his absence that year. In 2009, the relatively older child is around 5 years old and the younger sibling is in her/his early years. All outcomes are measured when both siblings are around 5 years old. Variable *younger sibling* is equal to 1 if the individual is the younger sibling and 0 if the individual is the relatively older child. Variable *hh with absence in 2009* is equal to 1 if the pair of siblings under comparison belong to a household which experienced the absence in 2009 but not in 2006 nor 2002, and 0 if the father was present in the first three rounds. Dependent variables in columns 1-2 are weight-for-age z-score and height-for-age z-score. Both anthropometric z-scores are based on the 2006 WHO child growth standards for children below 5 and 2007 WHO child growth standards for children aged 5-19 years (except for weight-for-age for children older than 10). Dependent variable in column 2 corresponds to the standardized Peabody Picture Vocabulary Test (PPVT) at the country level. For cognitive outcomes, India is excluded due to lack of data for younger siblings. All regressions are controlled for age, gender and birth order. Interactions between each control and each country are included.

Table E14: Nutritional and cognitive outcomes when 5, 8 and 12 years old, coefficients per country

VARIABLES	(1) Weight-for-age	(2) Height-for-age	(3) PPVT Z-Score
5 years old			
Ethiopia			
Younger sibling	0.571**	1.358*	0.0898
	(0.201)	(0.760)	(0.107)
HH w/absence in 2009 * younger sibling	0.0969	0.0951	-0.0881
	(0.207)	(0.193)	(0.0948)
Number of HH (treatment + control)	706	706	701
India			
Younger sibling	0.374	0.401	
	(0.316)	(0.550)	
HH w/absence in 2009 * younger sibling	0.191	0.0301	
	(0.180)	(0.110)	
Number of HH (treatment + control)	636	636	
Peru			
Younger sibling	0.136	1.024***	1.870***
	(0.105)	(0.182)	(0.260)
HH w/absence in 2009 * younger sibling	-0.0680	-0.0106	-0.123
	(0.195)	(0.191)	(0.0808)
Number of HH (treatment + control)	496	496	491
Vietnam			
Younger sibling	-0.544*	-0.269	0.200
	(0.281)	(0.320)	(0.182)
HH w/absence in 2009 * younger sibling	-0.338*	0.341	-0.0285
	(0.184)	(0.260)	(0.163)
Number of HH (treatment + control)	374	374	362
8 years old			
Ethiopia			
Younger sibling	0.457	0.841**	-2.370***
	(0.352)	(0.354)	(0.515)
HH w/absence in 2006 * younger sibling	0.0926	0.0484	0.0305
	(0.155)	(0.190)	(0.230)
Number of HH (treatment + control)	706	706	702
India			
Younger sibling	0.374	0.860***	
	(0.267)	(0.276)	
HH w/absence in 2006 * younger sibling	0.0481	0.0267	
	(0.159)	(0.0991)	
Number of HH (treatment + control)	636	636	
Peru			
Younger sibling	0.740	0.187	0.372*
	(0.480)	(0.287)	(0.184)
HH w/absence in 2006 * younger sibling	0.0594	0.0351	-0.0642
	(0.172)	(0.192)	(0.0815)
Number of HH (treatment + control)	496	496	495
Vietnam			
Younger sibling	1.400***	1.151***	-2.160***
	(0.453)	(0.374)	(0.442)
HH w/absence in 2006 * younger sibling	-0.189	0.123	-0.159
	(0.151)	(0.168)	(0.177)
Number of HH (treatment + control)	373	374	371
12 years old			
Ethiopia			
Younger sibling		-0.791	0.435***
		(0.854)	(0.0606)
HH w/absence in 2009 * younger sibling		0.0376	0.0923
		(0.153)	(0.0810)
Number of HH(treatment + control)		706	613
India			
Younger sibling		0.398***	
		(0.138)	
HH w/absence in 2009 * younger sibling		0.0511	
		(0.151)	
Number of HH(treatment + control)		635	
Peru			
Younger sibling		0.956	-1.179***
		(0.596)	(0.120)
HH w/absence in 2009 * younger sibling		0.109	-0.0335
		(0.216)	(0.0998)
Number of HH (treatment + control)		496	496
Vietnam			
Younger sibling		-0.283	-0.170
		(0.714)	(0.110)
HH w/absence in 2009 * younger sibling		0.495*	0.00359
		(0.259)	(0.115)
Number of HH (treatment + control)		374	373

Notes: Standard errors in parentheses, clustered at the sentinel site level. *** p<0.01, ** p<0.05, * p<0.1. Estimates are per country and outcomes are measured when aged 5, 8 and 12. Each regression includes a household fixed effect. Variable *younger sibling* is equal to 1 if the individual is the younger sibling and 0 if the individual is the relatively older child. Variable *hh with absence in 2009* is equal to 1 if the pair of siblings under comparison belong to a household which experienced the absence in 2009 but not in 2002 nor 2006, and 0 if the father was present in neither of those rounds. Dependent variables in column 1 and 2 are weight-for-age z-score and height-for-age z-score. Both anthropometric z-scores are based on the 2006 WHO child growth standards for children below 5 and 2007 WHO child growth standards for children aged 5-19 years (except for weight-for-age for children older than 10). Dependent variable in column 3 correspond to the standardized score of the Peabody Picture Vocabulary Test (PPVT) at the country level. For this outcome, India is excluded due to lack of data for younger siblings. All regressions are controlled for age, gender and birth order.

Table E15: Socio-emotional outcomes when 8 and 12 years old between younger and older siblings among households with and without absence in 2009

	Agency Index			Pride Index		
VARIABLES	(1)	(2)	(3)	(1)	(2)	(3)
8 years old						
HH w/absence in 2009 * younger sibling (Ethiopia)	0.757***	0.647***	0.487**	-0.156	0.338	0.369
	(0.233)	(0.232)	(0.198)	(0.244)	(0.357)	(0.243)
	[0.005]	[0.013]	[0.026]	[0.537]	[0.365]	[0.153]
HH w/absence in 2009 * younger sibling * India	-0.711	-0.605*	-0.546	-0.174	-0.315	-0.378
	(0.748)	(0.312)	(0.331)	(0.451)	(0.414)	(0.332)
	[0.949]	[0.843]	[0.828]	[0.402]	[0.914]	[0.971]
HH w/absence in 2009 * younger sibling * Peru	-0.686*	-0.479	-0.363	0.0417	-0.363	-0.257
	(0.390)	(0.344)	(0.330)	(0.294)	(0.392)	(0.291)
	[0.826]	[0.523]	[0.651]	[0.499]	[0.881]	[0.502]
HH w/absence in 2009 * younger sibling * Vietnam	-0.473	-0.609	-0.541	0.638	-0.367	-0.485
	(0.541)	(0.389)	(0.464)	(0.472)	(0.507)	(0.349)
	[0.576]	[0.906]	[0.902]	[0.257]	[0.938]	[0.654]
Observations	1,894	3,380	3,808	1,239	2,764	4,032
R-squared	0.031	0.023	0.022	0.106	0.079	0.083
Number of HH	954	1,709	1,924	624	1,398	2,036
Observations control	884	1,565	1,768	573	1,275	1,867
Observations treatment	70	144	156	51	123	169
P-value younger sibling	.063	.056	.026	.200	.151	.001
P-value HH w/absence in 2009 * younger sibling	.311	.223	.331	.481	.829	.534
12 years old						
HH w/absence in 2009 * younger sibling (Ethiopia)	-0.0114	0.0644	0.0667	0.0280	0.0251	0.0136
	(0.251)	(0.228)	(0.223)	(0.292)	(0.325)	(0.233)
	[0.965]	[0.784]	[0.773]	[0.926]	[0.940]	[0.955]
HH w/absence in 2009 * younger sibling * India	0.0672	-0.145	-0.0471	0.0234	-0.0741	-0.00547
	(0.396)	(0.274)	(0.262)	(0.393)	(0.373)	(0.282)
	[0.860]	[0.610]	[0.890]	[0.850]	[0.795]	[0.961]
HH w/absence in 2009 * younger sibling * Peru	-0.189	-0.258	-0.260	0.105	0.0591	0.0747
	(0.346)	(0.321)	(0.316)	(0.325)	(0.357)	(0.268)
	[0.419]	[0.412]	[0.407]	[0.377]	[0.582]	[0.522]
HH w/absence in 2009 * younger sibling * Vietnam	0.626	0.355	0.609*	-0.0672	-0.0373	-0.0869
	(0.499)	(0.368)	(0.362)	(0.437)	(0.402)	(0.336)
	[0.180]	[0.173]	[0.032]	[0.907]	[0.960]	[0.770]
Observations	3,244	4,145	4,189	2,581	3,419	4,220
R-squared	0.024	0.026	0.023	0.020	0.016	0.015
Number of HH	1,632	2,088	2,110	1,301	1,726	2,127
Observations control	1,502	1,917	1,939	1,191	1,576	1,954
Observations treatment	130	171	171	110	150	173
P-value younger sibling	.094	.094	.011	.000	.000	.001
P-value HH w/absence in 2009 * younger sibling	.433	.371	.118	.958	.950	.940

Notes: Standard errors in parentheses, clustered at the sentinel site level. *** p<0.01, ** p<0.05, * p<0.1. Each regression includes a household fixed effect and presents, by country, the gap in the corresponding outcome between a child and his/her younger sibling in households that experienced the absence of the father in 2009 versus the gap in household that did not experienced his absence that year. While the upper panel exhibits outcomes when both siblings were 8 years old, the bottom one shows them when they are 12 years old. The reference country corresponds to Ethiopia. Dependent variables in column 1-3 are the following subindices: (1) agency index including all questions, (2) agency index excluding those related to work and (3) agency index excluding those related to work and school. Dependent variables in column 4-6 are the following subindices: (1) pride index including all questions, (2) pride index excluding those related to work and (3) pride index excluding those related to work and school. These indices were calculated as explained in section 3. All regressions are controlled for age, gender and birth order. Interactions between each control and each country are included.

G Mechanisms

G.1 Income effect

Absence in 2009

Table F16 shows the evolution of food, non-food and total expenditures between the previous round (2006) and the one in which the father was absent in treated households (2009).[28] The estimate in the first line represents how expenditures change between households that experienced the absence of the father in 2009 versus those that did not at baseline, before the absence happened. Expenditures do not seem to differ on average when the father was present in both types of households.

[28]This analysis could not be done to the sample with absence in 2006 as no information on expenditures is available in the first round.

Once the father has left, expenditures evolved less on household experiencing absence with respect to the control households, as shown by the third estimate (the interaction term) where the difference is statistically significant for food expenditures at the 10% level. As an exploratory exercise, I compare the response of whether the household is economically the same, worse or better in 2006 versus in 2002. Table F17 shows that households experiencing absence in 2009 already had a higher likelihood of have a worse than before, even pre-departure of the father. This result somewhat weakens the argument that there were common trends before paternal absence for the mentioned sample. Further details on this issue will be exposed in section 9.

Table F16: Evolution of food, non-food and total expenditures pre and post paternal absence in 2009

VARIABLES	(1) Food expenditure	(2) Non-food expenditure	(3) Total expenditure
HH w/absence in 2009	0.0313	-0.0405	-0.0156
	(0.0773)	(0.0805)	(0.0809)
yr 2009	0.153***	0.106***	0.141***
	(0.0480)	(0.0318)	(0.0392)
HH w/absence in 2009 * yr 2009	-0.129*	-0.0356	-0.0816
	(0.0737)	(0.0766)	(0.0693)
India	0.0197	-0.0107	0.00332
	(0.118)	(0.109)	(0.118)
Peru	-0.00392	-0.0127	-0.00887
	(0.133)	(0.135)	(0.143)
Vietnam	0.00447	-0.000152	0.00323
	(0.174)	(0.158)	(0.179)
Constant	-0.0794	-0.0461	-0.0689
	(0.0940)	(0.0814)	(0.0868)
Observations	4,424	4,424	4,424
R-squared	0.005	0.003	0.005
Observations Control	2,030	2,030	2,030
Observations Treatment	182	182	182

Notes: Standard errors in parentheses, clustered at the sentinel site level. *** p<0.01, ** p<0.05, * p<0.1. No household fixed effects were included. Variable *yr 2009* is equal to 1 if the outcome corresponds to the year 2009 and 0 if it corresponds to 2006. Variable *hh w/absence in 2009* is equal to 1 if the father was absent in 2009 (but not in 2006 nor 2002) and 0 if present the three rounds. Food, non-food and total expenditures are reported monthly expenditures per capita in real 2006 local currency. In order to make values comparable, they were standardized at the country level. Negative values were reported as missing.

Table F17: Change in economic situation pre-absence of the father (2006 vs. 2002) in households with absence in 2009

VARIABLES	(1) Same	(2) Worse	(3) Better
HH w/absence in 2009	-0.0341	0.0693*	-0.0351
	(0.0368)	(0.0361)	(0.0343)
India	0.0937**	-0.0119	-0.0818**
	(0.0418)	(0.0311)	(0.0382)
Peru	0.289***	-0.144***	-0.145***
	(0.0546)	(0.0314)	(0.0469)
Vietnam	0.0625	-0.158***	0.0951
	(0.0656)	(0.0304)	(0.0668)
Constant	0.322***	0.223***	0.455***
	(0.0364)	(0.0250)	(0.0348)
Observations	2,212	2,212	2,212
R-squared	0.047	0.038	0.028
Observations Control	2,030	2,030	2,030
Mean Control	.422	.161	.417
Observations Treatment	182	182	182
Mean Treatment	.418	.225	.357

Notes: Standard errors in parentheses, clustered at the sentinel site level. *** p<0.01, ** p<0.05, * p<0.1. Variables are constructed from the following question in round 2 (2006): "During the period (current/four year ago), how would you describe the household you were living in?" (1) very rich (2) rich (3) comfortable-manage to get by (4) struggle-never have quite enough (5) poor (6) destitute. Variable same is equal to 1 if the situation has no changed and 0 otherwise. Variable worse is equal to 1 if current situation reported is higher (as the higher the poorer) than the one four years ago and 0 otherwise. Variable better is equal to 1 if current situation reported is lower than the one four years ago and 0 otherwise.

G.2 Family structure

Absence in 2006 and 2009

As I did for households where absence occurs in 2006, I perform the same analysis for households where absence occurs in 2009. Column 2 shows that treated households have the same probability of belonging to a nuclear family than those in the control group.

Table F18: Probability of belonging to a nuclear family

VARIABLES	(1) Absence in 2006	(2) Absence in 2009
HH w/absence (Ethiopia)	0.125**	-0.0224
	(0.0604)	(0.0795)
HH w/absence * India	-0.130	0.103
	(0.110)	(0.0979)
HH w/absence * Peru	-0.203*	-0.0193
	(0.120)	(0.106)
HH w/absence * Vietnam	0.00736	0.0337
	(0.126)	(0.101)
India	-0.327***	-0.337***
	(0.0361)	(0.0383)
Peru	-0.0235	-0.0201
	(0.0283)	(0.0288)
Vietnam	-0.162***	-0.164***
	(0.0318)	(0.0321)
Constant (Ethiopia)	0.697***	0.698***
	(0.0167)	(0.0182)
Observations	2,365	2,212
R-squared	0.083	0.080
Observations Control HH	2212	2030
Mean Control HH	.57	.57
Observations Treatment HH	153	182
Mean Treatment HH	.64	.56

Notes: This table reports the difference in the probability of belonging to a nuclear family between households that experienced absence either in 2006 (column 1) or 2009 (column 2) and the control group. Estimates are interacted with each country being Ethiopia the baseline. Nuclear family is defined as a household in which no other adult member apart from the father and the mother lives at home at baseline (2002). I consider adult any member older than 14.

Absence in 2009

Table F19 presents how the number of adults –excluding parents– changes over time for treated and untreated households. As estimates in the last panel show, the number of other members actually increased after the departure and it is statistically significant. The upper panel shows a placebo test for households that experienced absence in 2009 but before departure occurred (pre-treatment trends). No statistically significant difference in trends can be found before the absence occurs.

Table F19: Absence in 2009: Change in the number of adults over time

VARIABLES	(1) Number of adults
Absence in 2009	
Before departure	
HH w/absence in 2009	-0.0455
	(0.126)
year 2006	-0.441***
	(0.0499)
HH w/absence in 2009 * yr 2006	-0.103
	(0.0943)
Mean HH w/no absence at year 2002	1.05
Mean HH w/absence at year 2002	1.09
After departure	
HH w/absence in 2009	-0.0930
	(0.0819)
year 2009	-0.166***
	(0.0214)
HH w/absence in 2009 * yr 2009	0.320***
	(0.0878)
Mean HH w/no absence at year 2006	.78
Mean HH w/absence at year 2006	.73

Notes: Standard errors in parentheses, clustered at the sentinel site level. *** p<0.01, ** p<0.05, * p<0.1. Adult is defined as other members excluding parents, older than 14 years old by 2002. No household fixed effects were included. Each regression is controlled for country dummies.

Table F20: Absence in 2006: results including interaction with family structure

VARIABLES	(1) Weight-for-age	(2) Height-for-age	(3) PPVT Z-Score
5 years old			
Younger sibling	0.117	0.366	0.436**
	(0.255)	(0.343)	(0.206)
Younger sibling * nuclear family	-0.00900	0.0915	0.107***
	(0.0574)	(0.0664)	(0.0349)
HH w/absence in 2006 * younger sibling	-0.171	-0.202	0.234
	(0.150)	(0.138)	(0.143)
HH w/absence in 2006 * younger sibling * nuclear family	-0.0235	-0.00475	-0.159
	(0.198)	(0.228)	(0.132)
Observations control	2145	2122	1219
Observations control- nuclear families	1227	1210	800
Observations treatment	149	148	95
Observations treatment- nuclear families	96	96	68
8 years old			
Younger sibling	0.631***	0.804***	-1.737***
	(0.181)	(0.189)	(0.388)
Younger sibling * nuclear family	0.0283	0.104**	0.178**
	(0.0651)	(0.0516)	(0.0679)
HH w/absence in 2006 * younger sibling	-0.284	-0.152	0.417**
	(0.173)	(0.118)	(0.168)
HH w/absence in 2006 * younger sibling * nuclear family	0.235	-0.0222	-0.585**
	(0.240)	(0.144)	(0.257)
Observations control	1700	2132	1347
Observations control- nuclear families	968	1216	876
Observations treatment	106	149	109
Observations treatment- nuclear families	71	96	77
12 years old			
Younger sibling		0.113	-0.256
		(0.220)	(0.156)
Younger sibling * nuclear family		0.0769	0.0653*
		(0.0640)	(0.0326)
HH w/absence in 2006 * younger sibling		-0.195	0.0722
		(0.131)	(0.105)
HH w/absence in 2006 * younger sibling * nuclear family		0.0273	-0.0968
		(0.175)	(0.120)
Observations control		2117	1415
Observations control- nuclear families		1211	920
Observations treatment		148	110
Observations treatment- nuclear families		96	78
Includes India		Yes	No

Notes: Standard errors in parentheses, clustered at the sentinel site level. *** p<0.01, ** p<0.05, * p<0.1. The table reports the main results interacted with dummy variable *nuclear family* which is equal to 1 if the household is a nuclear family at baseline and 0 otherwise. Nuclear family is defined as a household in which no other adult member apart from the father and the mother lives at home at baseline (2002). I consider adult any member older than 14. Outcomes are measure when siblings are 5 (upper panel), 8 (middle panel) and 12 (bottom panel) years old. Variable *younger sibling* is equal to 1 if the individual is the younger sibling and 0 if the individual is the relatively older child. Variable *hh with absence in 2006* is equal to 1 if the pair of siblings under comparison belong to a household which experienced the absence in 2006 but not in 2002, and 0 if the father was present in both rounds. Dependent variables in column 1 and 2 are weight-for-age z-score and height-for-age z-score. Both anthropometric z-scores are based on the 2006 WHO child growth standards for children below 5 and 2007 WHO child growth standards for children aged 5-19 years (except for weight-for-age for children older than 10). Dependent variable in column 3 corresponds to the standardized score of the Peabody Picture Vocabulary Test (PPVT) at the country level. For this outcome, India is excluded due to lack of data for younger siblings. All regressions are controlled for age, gender and birth order. Interactions between each control and each country are included.

Table F21: Mother is absent in 2002 and 2006

	Mother absent in 2002		Mother absent in 2006	
	Control HH	Treated HH	Control HH	Treated HH
Ethiopia	2	0	3	1
Proportion	0.00	0.00	0.00	0.02
India	6	0	5	12
Proportion	0.01	0.00	0.01	0.36
Peru	0	0	4	2
Proportion	0.00	0.00	0.01	0.05
Vietnam	0	0	3	4
Proportion	0.00	0.00	0.01	0.15

Notes: The table shows the absolute number and proportion of cases of maternal absence in 2002 (first two columns and 2006 (next two columns), by type of households (control and treated).

Table F22: Mother is absent in 2002 and 2006

VARIABLES	(1) Weight-for-age+	(2) Height-for-age	(3) PPVT Z-Score
5 years old			
Younger sibling	0.114	0.411	0.540***
	(0.255)	(0.347)	(0.188)
Younger sibling * mother absent in 2006	-0.112	-0.389	-0.170
	(0.265)	(0.356)	(0.198)
HH w/absence in 2006 * younger sibling	-0.197**	-0.232*	0.137
	(0.0954)	(0.132)	(0.107)
HH w/absence in 2006 * younger sibling * mother absent in 2006	0.188	0.630	0.380
	(0.395)	(0.429)	(0.299)
Observations Control	2145	2122	1219
Mean Control	-1.34	-1.64	-.89
Observations Treatment	149	148	95
Mean Treatment	-1.1	-1.37	-.97
8 years old			
Younger sibling	0.651***	0.856***	-1.540***
	(0.172)	(0.179)	(0.333)
Younger sibling * mother absent in 2006	-0.0713	0.0361	-0.834*
	(0.387)	(0.415)	(0.476)
HH w/absence in 2006 * younger sibling	-0.106	-0.142*	0.0119
	(0.0935)	(0.0763)	(0.171)
HH w/absence in 2006 * younger sibling * mother absent in 2006	-0.128	-0.187	0.438
	(0.557)	(0.452)	(0.579)
Observations Control	1700	2132	1347
Mean Control	-1.45	-1.38	.78
Observations Treatment	106	149	109
Mean Treatment	-1.19	-1.21	.87
12 years old			
Younger sibling		0.164	-0.247
		(0.226)	(0.151)
Younger sibling * mother absent in 2006		-0.332	-0.0222
		(0.372)	(0.208)
HH w/absence in 2006 * younger sibling		-0.133	0.0124
		(0.0886)	(0.0418)
HH w/absence in 2006 * younger sibling * mother absent in 2006		-0.00978	-0.0829
		(0.444)	(0.319)
Observations Control		2117	1415
Mean Control		-1.44	.28
Observations Treatment		148	110
Mean Treatment		-1.21	.32
All years together			
Younger sibling	0.206	0.429*	-0.151
	(0.142)	(0.231)	(0.135)
Younger sibling * mother absent in 2006	0.0482	-0.159	-0.220
	(0.372)	(0.270)	(0.201)
HH w/absence in 2006 * younger sibling	-0.130	-0.204**	0.118
	(0.0917)	(0.0824)	(0.0745)
HH w/absence in 2006 * younger sibling * mother absent in 2006	-0.215	0.0913	-0.194
	(0.551)	(0.347)	(0.324)
Includes India	Yes	Yes	No
Age polynomial	No	No	Yes

Notes: +All years together excludes period when children are 12 years old.
Standard errors in parentheses, clustered at the sentinel site level. *** p<0.01, ** p<0.05, * p<0.1. The table reports the main results interacted with dummy variable *mother absent in 2006* which is equal to 1 if the household experienced maternal absence in 2006 and and 0 otherwise. Outcomes corresponds to ages 5, 8, 12 and all three years pooled together. Variable *younger sibling* is equal to 1 if the individual is the younger sibling and 0 if the individual is the relatively older child. Variable *hh with absence in 2006* is equal to 1 if the pair of siblings under comparison belong to a household which experienced the absence in 2006 but not in 2002, and 0 if the father was present in both rounds. Dependent variables in column 1 and 2 are weight-for-age z-score and height-for-age z-score. Both anthropometric z-scores are based on the 2006 WHO child growth standards for children below 5 and 2007 WHO child growth standards for children aged 5-19 years (except for weight-for-age for children older than 10). Dependent variable in column 3 corresponds to the standardized score of the Peabody Picture Vocabulary Test (PPVT) at the country level. For this outcome, India is excluded due to lack of data for younger siblings. All regressions are controlled for age, gender and birth order. Interactions between each control and each country are included.

G.3 Enrollment to education

Table F23: Likelihood of spending time at school when 5, 8 and 12 years old between younger and older siblings among households with and without and without absence in 2006- Excluding India

VARIABLES	5 years old (1) spent time at school	8 years old (2) spent time at school	12 years old (3) spent time at school
Younger sibling	0.277	0.0516	-0.00240
	(0.219)	(0.0963)	(0.0807)
HH w/absence in 2006 * younger sibling	0.119*	-0.0486	-0.00284
	(0.0643)	(0.0355)	(0.0199)
Observations	1,878	3,256	3,178
R-squared	0.081	0.031	0.027
Number of HH	939	1,628	1,592
Observations Control	868	1514	1480
Mean Control	.58	.89	.96
Observations Treatment	71	114	112
Mean Treatment	.49	.93	.96

Notes: Standard errors in parentheses, clustered at the sentinel site level. *** p<0.01, ** p<0.05, *p<0.1. Question on time use is asked for all children in the household aged 5 to 17. Variable Spent time in school is 1 if the child has spent at least 1 hour at school and comes from the following question: "Typically how many hours did each child in the household (aged between 5 and 17yrs) spend on the following activities during a typical day (from Monday to Friday) in the last week?: At school (including travelling time to school). ". All regressions include a household fixed effect and are controlled for age, gender and birth order. Interactions between each control and each country are included. India has been excluded from the analysis.

G.4 Temporary versus more permanent paternal absence

Table F24: Absence in 2006: results including interaction with absence in 2009 (next round), all years together

VARIABLES	(1) Weight-for-age+	(2) Height-for-age	(3) PPVT Z-Score
5 years old			
Younger sibling	0.114	0.402	0.490**
	(0.256)	(0.343)	(0.205)
younger sibling * absence in 2009	0.0269	0.0672	-0.0875
	(0.105)	(0.0872)	(0.0568)
HH w/absence in 2006 * younger sibling	-0.238*	-0.189	0.106
	(0.124)	(0.186)	(0.105)
HH w/absence in 2006 * younger sibling * absence in 2009	0.0994	-0.0788	0.120
	(0.235)	(0.289)	(0.120)
Observations control with absence in 2009	173	170	88
Observations treatment with absence in 2009	63	63	41
All years together			
Younger sibling	0.206	0.418*	-0.182
	(0.140)	(0.229)	(0.144)
Younger sibling * absence in 2009	0.0580	0.0996	-0.0354
	(0.0882)	(0.0755)	(0.0564)
HH w/absence in 2006 * younger sibling	-0.251*	-0.292**	0.0909
	(0.133)	(0.128)	(0.0620)
HH w/absence in 2006 * younger sibling * absence in 2009	0.253	0.106	0.00231
	(0.214)	(0.209)	(0.124)
Observations	6,823	13,194	7,163
R-squared	0.053	0.074	0.535
Observations control with absence in 2009	260	489	252
Observations treatment with absence in 2009	70	186	114
Includes India	Yes	Yes	No
Age polynomial	No	No	Yes

Notes: +All years together excludes period when children are 12 years old.
Standard errors in parentheses, clustered at the sentinel site level. *** p<0.01, ** p<0.05, * p<0.1. The table reports the main results interacted with dummy variable *absence in 2009* which is equal to 1 if the household experienced absence in 2009 and and 0 otherwise. Outcomes corresponds to ages 5, 8 and 12 pooled together. Variable *younger sibling* is equal to 1 if the individual is the younger sibling and 0 if the individual is the relatively older child. Variable *hh with absence in 2006* is equal to 1 if the pair of siblings under comparison belong to a household which experienced the absence in 2006 but not in 2002, and 0 if the father was present in both rounds. Dependent variables in column 1 and 2 are weight-for-age z-score and height-for-age z-score. Both anthropometric z-scores are based on the 2006 WHO child growth standards for children below 5 and 2007 WHO child growth standards for children aged 5-19 years (except for weight-for-age for children older than 10). Dependent variable in column 3 corresponds to the standardized score of the Peabody Picture Vocabulary Test (PPVT) at the country level. For this outcome, India is excluded due to lack of data for younger siblings. All regressions are controlled for age, gender and birth order. Interactions between each control and each country are included.

H Robustness checks

H.1 Controlling for negative shocks

Table G25: Absence in 2006: controlling for negative shocks in 2002, all years together

VARIABLES	(1) Weight-for-age^{+}	(2) Height-for-age	(3) PPVT Z-Score
Younger sibling	0.205	0.393*	-0.190
	(0.139)	(0.229)	(0.139)
Younger sibling * negative shock in 2002	-0.00757	0.0805	0.0814**
	(0.0533)	(0.0512)	(0.0355)
HH w/absence in 2006 * younger sibling	0.0786	-0.0693	0.202
	(0.159)	(0.120)	(0.131)
HH w/absence in 2006 * younger sibling * negative shock in 2002	-0.345*	-0.244*	-0.162
	(0.198)	(0.136)	(0.124)
Observations Control	3214	6162	3318
Mean Control	-1.42	-1.48	.07
Observations Treatment	198	435	267
Mean Treatment	-1.19	-1.26	.06
Household Fixed Effect	Yes	Yes	Yes
Includes India	Yes	Yes	Yes
Age polynomial	No	No	No

Notes: $^{+}$All years together excludes period when children are 12 years old.
Standard errors in parentheses, clustered at the sentinel site level. *** p<0.01, ** p<0.05, * p<0.1. The table reports the main results interacted with dummy variable *negative shocks 2002* which is equal to 1 if the household reports to had experienced at least one of a listed shocks in 2002 and 0 otherwise. Outcomes corresponds to ages 5, 8 and 12 pooled together. Negative shocks include a wide variety of events (economic, environmental, changes in family structure, agricultural socks, etc). Variable *younger sibling* is equal to 1 if the individual is the younger sibling and 0 if the individual is the relatively older child. Variable *hh with absence in 2006* is equal to 1 if the pair of siblings under comparison belong to a household which experienced the absence in 2006 but not in 2002, and 0 if the father was present in both rounds. Dependent variables in column 1 and 2 are weight-for-age z-score and height-for-age z-score. Both anthropometric z-scores are based on the 2006 WHO child growth standards for children below 5 and 2007 WHO child growth standards for children aged 5-19 years (except for weight-for-age for children older than 10). Dependent variable in column 3 corresponds to the standardized score of the Peabody Picture Vocabulary Test (PPVT) at the country level. For this outcome, India is excluded due to lack of data for younger siblings. All regressions are controlled for age, gender and birth order. Interactions between each control and each country are included.

Table G26: Absence in 2006: controlling for negative shocks in 2006, all years together

VARIABLES	(1) Weight-for-age^{+}	(2) Height-for-age	(3) PPVT Z-Score
Younger sibling	0.159	0.376	-0.152
	(0.146)	(0.237)	(0.134)
Younger sibling * bad event in 2006	0.0493	0.0596	-0.00972
	(0.0568)	(0.0611)	(0.0478)
HH w/absence in 2006 * younger sibling	-0.000851	-0.187	0.0946
	(0.214)	(0.235)	(0.130)
HH w/absence in 2006 * younger sibling * bad event in 2006	-0.173	-0.0312	0.00768
	(0.253)	(0.240)	(0.146)
Observations Control	3214	6162	3318
Mean Control	-1.42	-1.48	.07
Observations Treatment	198	435	267
Mean Treatment	-1.19	-1.26	.06
Household Fixed Effect	Yes	Yes	Yes
Includes India	Yes	Yes	Yes
Age polynomial	No	No	No

Notes: $^{+}$All years together excludes period when children are 12 years old.
Standard errors in parentheses, clustered at the sentinel site level. *** p<0.01, ** p<0.05, * p<0.1. The table reports the main results interacted with dummy variable *negative shocks 2006* which is equal to 1 if the household reports to had experienced at least one of a listed shocks in 2006 and 0 otherwise. Negative shocks include a wide variety of events (economic, environmental, changes in family structure, agricultural socks, etc). Variable *younger sibling* is equal to 1 if the individual is the younger sibling and 0 if the individual is the relatively older child. Variable *hh with absence in 2006* is equal to 1 if the pair of siblings under comparison belong to a household which experienced the absence in 2006 but not in 2002, and 0 if the father was present in both rounds. Dependent variables in column 1 and 2 are weight-for-age z-score and height-for-age z-score. Both anthropometric z-scores are based on the 2006 WHO child growth standards for children below 5 and 2007 WHO child growth standards for children aged 5-19 years (except for weight-for-age for children older than 10). Dependent variable in column 3 corresponds to the standardized score of the Peabody Picture Vocabulary Test (PPVT) at the country level. For this outcome, India is excluded due to lack of data for younger siblings. All regressions are controlled for age, gender and birth order. Interactions between each control and each country are included.

H.2 Alternative controls

Table G27: Absence in 2006: results with alternative controls

	Weight-for-age+			Height-for-age			PPVT Z-score		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	All	No future absence	Future absence	All	No future absence	Future absence	All	No future absence	Future absence
5 years old									
HH w/absence in 2006 * younger sibling	-0.187**	-0.185**	-0.192*	-0.199	-0.196	-0.238**	0.126	0.115	0.175*
	(0.0870)	(0.0882)	(0.100)	(0.120)	(0.127)	(0.119)	(0.0942)	(0.0935)	(0.103)
Observations Control	2,145	1,784	360	2,122	1,765	356	1,219	1,017	201
Mean Control	-1.34	-1.33	-1.37	-1.64	-1.63	-1.67	-.89	-.88	-.95
Observations Treatment	149	149	149	148	148	148	95	95	95
Mean Treatment	-1.1	-1.1	-1.1	-1.37	-1.37	-1.37	-.97	-.97	-.97
8 years old									
HH w/absence in 2006 * younger sibling	-0.125	-0.126	-0.122	-0.160**	-0.166**	-0.151	0.0138	0.0242	-0.0681
	(0.0924)	(0.0913)	(0.108)	(0.0775)	(0.0796)	(0.0907)	(0.154)	(0.155)	(0.157)
Observations Control	1,700	1,414	285	2,132	1,769	362	1,347	1,119	227
Mean Control	-1.45	-1.45	-1.42	-1.38	-1.37	-1.43	.78	.83	.51
Observations Treatment	106	106	106	149	149	149	109	109	109
Mean Treatment	-1.19	-1.19	-1.19	-1.21	-1.21	-1.21	.87	.87	.87
12 years old									
HH w/absence in 2006 * younger sibling				-0.173*	-0.164*	-0.236**	0.00669	0.00410	0.00482
				(0.0890)	(0.0906)	(0.108)	(0.0457)	(0.0475)	(0.0550)
Observations Control				2117	1760	357	1415	1177	237
Mean Control				-1.44	-1.43	-1.52	.28	.27	.32
Observations Treatment				148	148	148	110	110	110
Mean Treatment				-1.21	-1.21	-1.21	.32	.32	.32
All years together									
HH w/absence in 2006 * younger sibling	-0.145	-0.145	-0.146	-0.212***	-0.209**	-0.250***	0.0795	0.0808	0.0648
	(0.0910)	(0.0898)	(0.108)	(0.0792)	(0.0807)	(0.0905)	(0.0715)	(0.0718)	(0.0778)

Notes: +All years together excludes period when children are 12 years old.
Standard errors in parentheses, clustered at the sentinel site level. *** p<0.01, ** p<0.05, * p<0.1. The table reports the main results varying the definition of the control households. Columns 1, 4 and 7 keeps the original definition, fathers are present in 2002 and 2006 with no restriction afterward. Column 2, 5 and 8 impose that the father has also to be present in 2009, 2013 and 2016 for the control households. Column 3, 6 and 9 impose that the father can is absent either in 2009 or 2013 or 2016. No extra restrictions are imposed to the treated households given the small sample size. Variable *younger sibling* is equal to 1 if the individual is the younger sibling and 0 if the individual is the relatively older child. Variable *hh with absence in 2006* is equal to 1 if the pair of siblings under comparison belong to a household which experienced the absence in 2006 but not in 2002, and 0 depending on the definition for the control group. Dependent variables in column 1 and 2 are weight-for-age z-score and height-for-age z-score. Both anthropometric z-scores are based on the 2006 WHO child growth standards for children below 5 and 2007 WHO child growth standards for children aged 5-19 years (except for weight-for-age for children older than 10). Dependent variable in column 3 corresponds to the standardized score of the PPVT at the country level. For this outcome, India is excluded due to lack of data for younger siblings. All regressions are controlled for age, gender and birth order. Interactions between each control and each country are included.

Table G28: Absence in 2009: results with alternative controls

	Weight-for-age+			Height-for-age			PPVT Z-score		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	All	No future absence	Future absence	All	No future absence	Future absence	All	No future absence	Future absence
5 years old									
HH w/absence in 2009 * younger sibling	0.0265	0.0273	0.0178	0.0639	0.0638	0.0739	-0.0898	-0.102*	-0.0141
	(0.105)	(0.105)	(0.140)	(0.0873)	(0.0911)	(0.142)	(0.0572)	(0.0586)	(0.0921)
Observations Control	1972	1784	187	1952	1765	186	1131	1017	113
Mean Control	-1.34	-1.33	-1.42	-1.64	-1.63	-1.74	-.88	-.88	-.95
Observations Treatment	173	173	173	170	170	170	88	88	88
Mean Treatment	-1.31	-1.31	-1.31	-1.59	-1.59	-1.59	-.94	-.94	-.94
8 years old									
HH w/absence in 2009 * younger sibling	0.0389	0.0354	0.0907	0.0429	0.0357	0.0943	-0.0427	-0.0278	-0.180
	(0.0849)	(0.0839)	(0.127)	(0.0810)	(0.0789)	(0.136)	(0.0938)	(0.0936)	(0.135)
Observations Control	1556	1414	141	1957	1769	187	1243	1119	123
Mean Control	-1.46	-1.45	-1.48	-1.38	-1.37	-1.44	.8	.83	.49
Observations Treatment	144	144	144	175	175	175	104	104	104
Mean Treatment	-1.36	-1.36	-1.36	-1.42	-1.42	-1.42	.54	.54	.54
12 years old									
HH w/absence in 2009 * younger sibling				0.125	0.124	0.114	0.0125	0.0147	-0.00243
				(0.0952)	(0.0949)	(0.141)	(0.0583)	(0.0604)	(0.0701)
Observations Control				1941	1760	181	1307	1177	129
Mean Control				-1.44	-1.43	-1.55	.27	.27	.22
Observations Treatment				176	176	176	108	108	108
Mean Treatment				-1.48	-1.48	-1.48	.44	.44	.44
All years together									
HH w/absence in 2009 * younger sibling	0.0561	0.0536	0.0817	0.0979	0.0951	0.111	-0.0328	-0.0286	-0.0834
	(0.0885)	(0.0874)	(0.123)	(0.0757)	(0.0754)	(0.114)	(0.0551)	(0.0567)	(0.0611)

Notes: +All years together excludes period when children are 12 years old.
Standard errors in parentheses, clustered at the sentinel site level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. The table reports the main results varying the definition of the control households. Columns 1, 4 and 7 keeps the original definition, fathers are present in 2002, 2006 and 2009 with no restriction afterward. Column 2, 5 and 8 impose that the father has also to be present in 2013 and 2016 for the control households. Column 3, 6 and 9 impose that the father can is absent either 2013 or 2016. No extra restrictions are imposed to the treated households given the small sample size. Variable *younger sibling* is equal to 1 if the individual is the younger sibling and 0 if the individual is the relatively older child. Variable *hh with absence in 2009* is equal to 1 if the pair of siblings under comparison belong to a household which experienced the absence in 2009 but not in 2006 nor 2002, and 0 depending on the definition for the control group. Dependent variables in column 1 and 2 are weight-for-age z-score and height-for-age z-score. Both anthropometric z-scores are based on the 2006 WHO child growth standards for children below 5 and 2007 WHO child growth standards for children aged 5-19 years (except for weight-for-age for children older than 10). Dependent variable in column 3 corresponds to the standardized score of the PPVT at the country level. For this outcome, India is excluded due to lack of data for younger siblings. All regressions are controlled for age, gender and birth order. Interactions between each control and each country are included.