Essays on Labor Market Outcomes for Women in Developing Countries

Laine Rutledge

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2017

Reading Committee:

Rachel Heath, Chair

Brian Dillon

Judith Thornton

Program Authorized to Offer Degree:
Department of Economics

University of Washington

Abstract

Essays on Labor Market Outcomes for Women
in Developing Countries

Laine Rutledge

Chair of the Supervisory Committee:
Assistant Professor Rachel Heath
Economics

This dissertation focuses on labor market issues for women in developing countries. The first chapter investigates adult life outcomes of a child sponsorship program in six developing countries, with particular attention paid to differing outcomes for men and women. Results show positive impacts on total years of education, and primary, secondary, and university school completion. Particularly large impacts are seen for women in areas where baseline female education is low. The second chapter analyzes the return to child sponsorship, and how these returns vary between men and women. We find positive effects of child sponsorship on men's income, while the positive return to sponsorship for women is limited to increased income due to increased labor market participation. The third chapter studies labor market migration in Brazil. I use individual and firm fixed-effects to examine the return to moving for men and women. The addition of firm fixed effects does not greatly impact the return to moving for men but eliminates the positive returns measured for women when only using individual fixed-effects. This indicates that any promotion and geographical movement that women experience is not reflected in earnings.

TABLE OF CONTENTS

Page

LIST OF FIGURES

LIST OF TABLES

Table Number Page

ACKNOWLEDGMENTS

I would like to express my gratitude to my advisor, Professor Rachel Heath, and committee member Professor Brian Dillon for their support and guidance. Their vast knowledge, comments, suggestions, and patience were instrumental to my graduate school success. I would also like to thank committee members Professors Elaina Rose, Judith Thornton, and Darryl Holman for their input and encouragement.

In addition to my committee, I would like to thank Professors Bruce Wydick and Paul Glewwe. Their mentorship and collaborative spirit allowed me to grow as a researcher and I am indebted to them for their support throughout graduate school. I also express gratitude to Ian Schmutte, University of Georgia, and Jason Rivera for providing an English codebook for the RAIS data.

I would also like to thank Jennifer Meredith, Marie Guldin, Abigail Harley, and Jenny Ho for their unending support, feedback, and encouragement. I am also thankful for the support and cheer of John and Adele.

Finally, I want to thank my family; my mom, dad, Warren, Amanda, my brother, sister-in-law, and nephews; for their understanding, encouragement, and motivation.

DEDICATION

To Peter, I carry you with me always.

Chapter 1

DOES CHILD SPONSORSHIP WORK? A SIX-COUNTRY STUDY OF IMPACTS ON ADULT LIFE OUTCOMES

The published version of this paper can be found in *Journal of Political Economy* 121 (2): 393 - 436. University of Chicago Press.

Bruce Wydick[1], Paul Glewwe[2], and Laine Rutledge

Abstract

International child sponsorship is one of the leading forms of direct aid from households in wealthy countries to needy children in developing countries, where we estimate that 9.14 million children are currently supported through formal international sponsorship organizations. In a study involving original data collection on 10,144 individuals in six countries, we present estimated impacts on adult life outcomes from sponsorship through Compassion International, a leading child sponsorship organization. To generate counterfactuals for identifying program effects, we utilize an age-eligibility rule that was followed from 1980 to 1992 as the program was being introduced into villages in those countries. We find large and statistically significant impacts from child sponsorship on years of completed schooling, primary, secondary, and tertiary school completion, and on the probability and quality of

---

[1]Wydick, Professor, Department of Economics, University of San Francisco, 2130 Fulton Street, San Francisco, CA 94117-1080, e-mail: wydick@usfca.edu.

[2]Glewwe, Professor, Department of Applied Economics, University of Minnesota, 1994 Buford Ave, St. Paul, MN 55108, e-mail: pglewwe@umn.edu.

adult employment. We summarize early evidence which suggests that these impacts may be due in part to programming that raises the aspirations and self-expectations of impoverished children.

1.1 Introduction

For millions of households in wealthy countries, international child sponsorship represents the most intimate and direct form of involvement with the poor in developing countries. Sponsors typically give $25-40 per month to supplement an impoverished childs education and health expenses or to support programs in which the child participates. We estimate that currently private financial flows to internationally sponsored children exceed US$3 billion annually, yet no published research exists that has gauged these programs impacts on the life outcomes of sponsored children.

This paper examines whether children sponsored through Compassion International, a leading child sponsorship organization currently serving 1.3 million children in 26 countries, have improved life outcomes when they are adults. Data were collected on the life outcomes of 10,144 individuals over two years from six developing countries that are representative of the Compassion programs work worldwide: Bolivia, Guatemala, India, Kenya, the Philippines, and Uganda.

Identification of the programs impacts rests on three eligibility rules that Compassion used from 1980 to 1992, when those in our study were sponsored. These rules limited the number of children per household that could be sponsored, required sponsored children to be within walking distance of a project, and stipulated that only children below a given age were eligible for the program. Using several estimation strategies that harness these eligibility rules to construct counterfactuals for estimating causal effects, we find that the Compassion sponsorship program significantly increased total years of schooling and completion rates across all levels of schooling. Impacts are especially large for secondary school completion, which increased by 12-18 percentage points over an average baseline of 44.5%. Education impacts are particularly strong in the two African countries. We also find positive and significant impacts on the probability of adult employment and movement into white-collar jobs. In some respects, Compassion projects are similar to many government and international donor programs that promote education. Sponsors pay for childrens school tuition and uniforms, several nutritious meals per week, healthcare, and tutoring. What distinguishes Compassion projects from most government and international donor programs, and from some other child sponsorship programs, is that children spend at least eight hours per week in an intensive after-school program that emphasizes their spiritual, physical, and socio-emotional development. In the sample, the average duration of sponsorship was 9.3 years, so that by the end of their childhood, sponsored children have participated in about 4,000 hours of Compassion programming, including extra activities such as retreats and camps. A primary objective of this extended contact is to raise the childs self-esteem, aspirations, and self-expectations.

Recent work in economics suggests that internal constraints that reflect low aspirations and reference points may lead to poverty traps (Ray, 2006; Dalton et al., 2010; Bernard et al., 2011). After reporting program impacts on adult life outcomes, we present a short summary of evidence from three follow-up studies that collected data and carried out other research among 1,380 currently sponsored Compassion children in Bolivia, Kenya, and Indonesia. These studies find that sponsored children exhibit significantly higher levels of self-esteem, aspirations and self-expectations, and lower levels of hopelessness. While more work is needed to establish a causal link between aspirations during youth and adult life outcomes, a clearer understanding of this relationship may have important implications for the way in which practitioners approach development work. Two major empirical issues must be addressed to obtain unbiased estimates of the impact of this type of program. First, selection of children into a program may not be random; more needy children may have been chosen (as is directed in the Compassion operations manual), but it is also possible that parents chose the children whom they thought were most likely to succeed. Second, program impacts may spill over onto siblings or non-treated peers in a treated village.

This study uses three program eligibility rules to address these two estimation issues. As Compassion established new projects from 1980 to 1992 in each of the six countries, an age-eligibility rule stipulated that only children 12 years and younger (11 years and under in Uganda and Guatemala) at the time the project started were eligible for sponsorship. This arbitrary rule allows us to compare the adult life outcomes of formerly sponsored children relative to the life outcomes of their ineligible older siblings, who were age 13 or older when the program was started in their village. Moreover, to maximize the number of households benefiting from sponsorships, a second rule set an upper limit on the number of sponsored children per household; this number varied from one to three in the six countries. Finally, a third rule stipulated that, to be sponsored, a child had to reside within walking distance of the program center, which was usually interpreted as at most a thirty-minute walk. In practice, this meant that only children residing within the village where the program was located were chosen for sponsorship. Children from neighboring villages were excluded.

We use the first two rules to address the issue of endogenous child selection. Specifically, we use a vector of dummy variables that indicate the age of an individual when the Compassion program was introduced into that persons village, interacted with his or her sibling order relative to program rollout (oldest sibling 12 or younger when the program began, second-oldest sibling among those 12 or younger when the program began, etc.), as instruments to estimate the probability that an individual was a sponsored child. These instruments are strongly correlated with take-up since children meeting the age requirement were far more likely to be sponsored, and in practice the oldest age-eligible siblings were most likely to be sponsored. These instruments satisfy the exclusion restriction for instrumental variables because as Compassion programs were introduced in different years in different villages, the age of an individual at the time the program began in his or her village should not be related to his or her life outcomes except via the impact of being sponsored.

We use the third rule to address the issue of program spillovers; we collected data not only on sets of siblings from treated families, but also on sets of siblings from a random sample of non- treated households in Compassion villages and from a random sample of households in neighboring villages where children could not be sponsored due to the walking distance rule. Using the identifying assumptions that program spillovers affect neither older age-ineligible children in treated villages, nor any children in non-Compassion villages, we implement both OLS and IV difference-in- differences regressions that estimate direct effects on program participants and can also be used to calculate both intra-household and intra-village spillovers from the Compassion program.

These estimates yield large and statistically significant effects of child sponsorship on education, employment, and leadership outcomes. OLS and IV estimations, with and without the use of household fixed effects, find that child sponsorship resulted in 1.03 to 1.46 additional years of completed schooling for sponsored children over a baseline of 10.19 years for unsponsored children.[3] Impacts on primary schooling range from 4.0 to 7.7 percentage points over an untreated baseline of 88.7%. Impacts on secondary school completion are greater and highly significant, ranging from 11.6 to 16.5 percentage points over a baseline of 44.5%; accounting for marginally significant spillover effects pushes the figure somewhat higher, to 13.7 - 18.5 percentage points. Tertiary education point estimates of impact are smaller, from 2.1 to 3.6 percentage points, but these are realized over a small baseline of 4.3%. Child sponsorship also appears to be a great "equalizer." It's impacts on the educational outcomes are larger in those countries with lower baseline education outcomes, the two African countries, while impacts in Latin America and Asia are smaller, although still

[3] "Years of schooling" denotes highest grade attained. Grade repetition is common in many countries, so years of schooling can exceed grade attained; we have no data on repetition so we cannot account for repeated grades.

statistically significant. Similarly, in countries where baseline schooling is higher for boys, child sponsorship tends to have a bigger impact on girls; where it is higher for girls, it has a bigger impact on boys.

OLS estimates also find positive impacts on the probability of salaried employment (5.1-6.3 percentage points, over a 35.7% baseline) and white-collar work (6.5-6.7 percentage points, baseline of 18.5%). There is also evidence, albeit mixed, of increases in community and church leadership.[4]

Our results do not necessarily apply to all child sponsorship programs. While some of the other major child sponsorship organizations, such as Children International, use sponsor funding for the nurturing and development of individual children, other large programs, such as those operated by World Vision, Plan USA, and Save the Children, use funding given in the name of a sponsored child more broadly to create village-level public goods. The less-targeted nature of these programs renders potential impacts more diffuse, and thus more difficult to assess. The remainder of this paper is organized as follows. Section 2 reviews the most relevant previous studies. Section 3 explains the fieldwork and data collection. Section 4 describes the estimation strategy, and Section 5 presents estimates of the program impacts. Section 6 concludes and briefly discusses potential causal mechanisms through the impact of the program on childrens aspirations, summarizing results from three follow-up studies of currently sponsored children.

1.2 Existing Research and Literature

Given the number of individuals involved in child sponsorship relationships and the billions of dollars committed to them, it is surprising that almost no research exists that evaluates the impacts of these programs. One exception is Kremer et al. (2003), who use a randomized

[4] Results on adult life outcomes such as age of marriage, fertility, remittances, dwelling quality, and durable good ownership are found in our larger working paper (Wydick, Glewwe and Rutledge, 2013a).

experiment to analyze the impacts of a Dutch child-sponsorship program that funded new classroom construction and provided students a \$6 uniform and \$3.44 worth of textbooks. They find that even these relatively low-cost interventions induced student beneficiaries to attend school a half year longer, and to advance a third of a grade farther in formal education.

More generally, a growing literature attempts to find cost-effective ways to induce parents to invest more in their children's education. Researchers have studied many programs, including cash transfers, free meals, provision of school uniforms, deworming medicine, and free medical treatment.

Perhaps the best known and most frequently evaluated intervention is the PROGRESA (later renamed Oportunidades) conditional cash transfer (CCT) program. Implemented in 1997 in poor regions of Mexico, Oportunidades provides cash incentives for mothers to increase their childrens school attendance and obtain health care for younger children. It was initially implemented as a randomized trial to facilitate its evaluation by researchers. Impact evaluations have shown that Oportunidades led to higher school enrollment, lower grade repetition, lower dropout rates, and higher school reentry rates among dropouts (Behrman et al., 2005; Bank, 2009). Behrman et al. (2007) estimate that receiving Oportunidades cash transfers for 5.5 years increased grades completed by 0.8 to 1.0 years. Schultz (2004) estimates that Oportunidades increased formal schooling by 0.66 years (0.72 for girls, and 0.64 for boys). He also finds that Oportunidades raised enrollment by 3.4 percentage points, averaging over all children in grades one through eight, with much larger impacts in later grades, not only for girls (14.8 percentage points) but also for boys (6.5 percentage points). Bobonis and Finan (2009) find that enrollment rates in Oportunidades communities increased by 5 percentage points, even among those ineligible for the program.

CCT programs have also had positive impacts on education in other countries. For example, Barrera-Osorio et al. (2008) implemented a randomized experiment to evaluate the *Conditional Subsidies for School Attendance* program in Bogota, Colombia, finding that school attendance increased by 2.8 percentage points on average.

Other programs have funded various school "inputs," such as free or subsidized school meals, uniforms, textbooks, school construction, and teachers. Several focus on nutrition or health. Drèze and Kingdon (2001) find that providing a mid-day meal in India raised girls school attendance by 15 percentage points. Kremer and Vermeersch (2004) estimate that school attendance rose by 8.5 percentage points in Kenyan preschools that provided free meals, increasing attendance of current students and attracting new students who had never attended preschool. Handa and Peterman (2007) find that South African childrens educational attainment is strongly affected by their nutritional status. Glewwe and Miguel (2008) review the impact of health and nutrition on education outcomes.

Many randomized studies of education interventions have been conducted in Kenya. Evans et al. (2008) evaluate a program that selected Kenyan children by lottery to receive free school uniforms. They find that receiving a uniform reduced absenteeism by 39 percent, and by 64 percent for poorer students. In the same area of Kenya, Miguel and Kremer (2004) implemented a randomized de-worming intervention. This intervention not only decreased overall disease transmission but also reduced school absenteeism by 7 percentage points in the treatment schools. They also find positive spillover effects onto children who attended nearby schools that did not participate in the de-worming program. In a follow-up study of former participants ten years after the deworming experiment, Ozier (2011) finds increases in cognitive performance equal to 0.5 to 0.8 years of schooling. A randomized trial that provided sixth grade girls merit scholarships of about $20 for school fees and school supplies increased student attendance by 5 percentage points; surprisingly, it increased both girls' and boys' test scores (Kremer et al., 2009). Another study in Kenya provided incentives to teachers to improve their teaching. Glewwe et al. (2010) carried out an experiment that provided valuable prizes to teachers based on their students test scores. Despite the incentives, teacher attendance did not improve; instead teachers held additional prep sessions prior to the exams on which the incentives were based, which led to only short-term increases in test scores.

Methodologically, the empirical strategy of this paper is similar to that of Duflo (2001) in that it uses the ages of former students and geographic placement of a schooling treatment as instruments to identify program impacts. Using a method similar to that of Pitt et al. (1993), Duflo examines the impact of Indonesias rapid expansion of school construction from 1973 to 1979. She uses an individuals exposure to the program, as measured by the number of schools built in his or her region of birth, along with age at the time of program implementation, to identify impacts on education and wages. She finds that each new school constructed per 1,000 children led to a 0.12 to 0.19 increase in years of schooling. This implies an average increase of 0.25 to 0.40 years per child beneficiary (about two schools were built per 1000 children), which then resulted in a 3.0 to 5.4 percent increase in wages, suggesting an economic return to education of 6.8 to 10.6 percent. She also finds that those who benefited most were among the poorest.

1.3 Counting the Sponsored, Program Background, and Fieldwork

1.3.1 Counting the Sponsored

There have been no reliable figures on the number of internationally sponsored children worldwide; a preliminary task was to estimate this figure. Through internet searches in multiple languages and contact with industry personnel across countries, we tallied 207 organizations that appear to represent nearly all children sponsored through such organizations worldwide. Based on the sponsorship figures claimed by these organizations, we estimate that there are currently 9.14 million internationally sponsored children in the world.[5] Over

---

[5]Because the internet is so vital today for fundraising in the child sponsorship industry, for example posting pictures of children and providing other contact between potential sponsors and potentially sponsored children, we assumed that any such organization of significant scope must have an internet presence. This

90% of this total are sponsored through the ten largest sponsorship organizations. Table 1.1 contains basic information about these organizations, including years of operation, number of countries served, monthly sponsorship fees, and number of children sponsored. All are based in the United States and Europe, and two of the largest three are faith-based, as are four of the largest ten.

The total flow of child sponsorship funds to developing countries is non-trivial; indeed it is similar to amounts given by the U.S. Government for international assistance. Assuming an average monthly contribution of $30, funding for child sponsorship is about US$3.29 billion per year, excluding special gifts and travel to visit sponsored children. This is comparable to USAID budgets of $8.72 billion in 2012 for "Global Health and Child Survival," $2.92 billion for "Development Assistance" and the $1.12 billion earmarked for the Millennium Challenge Corporation (Office of Management and Budget, 2012).

International child sponsorship programs arose due to their usefulness as a marketing tool for mobilizing resources in rich countries to reduce poverty in poor countries. As the marketers of these programs have recognized for decades, contact with an individual child creates a commitment device to help donors contribute a fraction of their monthly income to alleviating child poverty in developing countries via a relationship with a *particular child* living in poverty. In this way, international child sponsorship programs mobilize resources by drawing on the psychological and moral instincts people possess to care for their own children. Even in difficult economic times, the commitment of donors to the well-being of "their child" is likely to exceed their commitment to a large, well-intentioned – yet relatively faceless – non-profit organization.[6] Thus even apart from issues of impact and cost-effectiveness, child

assumption is the basis for our calculation of 9.14 million sponsored children. If there are child sponsorship programs that do not use the internet, our 9.14 million figure would underestimate the actual number of children sponsored worldwide.

[6]There is at least anecdotal evidence of this: During the first year of the 2008-09 recession, when giving to most U.S. charities declined sharply, World Vision reported that the percentage of those who maintained

sponsorship programs may be among the most effective methods for mobilizing resources to benefit children in developing countries.

1.3.2 The Compassion Child Sponsorship Program

The worlds third largest child sponsorship program is Compassion International, a large, faith-based, nonprofit organization whose stated goal is to "release children from spiritual, economic, social, and physical poverty." Compassion staffs its projects locally, and foreign employees are rare. The projects rely on volunteers from local churches and other organizations to carry out its programming. The benefits sponsored children receive vary somewhat by country, and even within countries, and Compassions approach has evolved over time. Table 1.2 summarizes, for each country in this study, the benefits the sponsored children received while enrolled in the program. In Uganda and Kenya, and in three of the projects in Bolivia, Compassion operated student centers where sponsored children gathered on Saturday or after school on weekdays. Students participated in structured programs at these centers, receiving academic tutoring, spiritual instruction, healthcare, nutritious meals and school supplies. They also participated in a wide array of games and activities. In most of the projects with student centers, Compassion children also received school fee subsidies and school uniforms. Compassion typically sponsors children through secondary school, although a small number continue to the university level through its leadership development program.

In Guatemala and the Philippines, Compassion programs operated in (Protestant) Christian schools, where students would receive similar benefits, although tutoring was not generally an explicit component of sponsorship. In India and in one project in Bolivia, Compassion collaborated with government programs that gave parents direct cash payments conditional upon the sponsored child's continuation in school. However, these programs differed from

---

their monthly financial commitment to sponsored children showed no sign of decline during that period (Kennedy, 2009) standard CCT programs in that children received most of the benefits provided by the other Compassion programs, as well as individual nurturing and care via Compassion's partnership with local Protestant churches.

All projects provided basic healthcare benefits. They included regular physical examinations administered by local nurses and doctors at Compassion schools and student centers. Also included was a form of catastrophic health insurance paid through a separate fund operated by Compassion's headquarters in Colorado. If a Compassion child had a serious illness or needed surgery, this fund covered the full cost of the procedure and hospitalization. In the rare cases when children needed such care, this benefit was often reported by formerly sponsored children to be the greatest source of support offered by the program. Aside from this catastrophic insurance, however, all funds directed to children flow from their sponsors regular monthly contributions.

All children sponsored through Compassion write letters several times per year to their sponsor, and most receive correspondence from their sponsor (71.8% in our study). In addition, about once per year sponsors receive a picture of the child and updates from local Compassion staff on the child's progress in school. Most children (83.7%) also reported receiving birthday gifts from their sponsor. Sponsors can also travel on organized trips to visit their sponsored children and their families; while not uncommon, this was not the norm.

The survey included an open-ended question asking formerly sponsored children which component of the Compassion program had been most beneficial to them. The most common answer was educational support (38.5%). (Within this category, payment of school fees and tutoring were cited most often, and almost equally.) The second-most common response was spiritual and character development (29.4%), followed by economic aid (9.5% – a figure that was no higher in the two countries where parents received direct cash payments), healthcare benefits (2.8%), and gifts received from sponsors (0.8%).

In Compassion projects, selection of children for sponsorship is done locally. Compassion instructs its staff to work with local community members to select children using six criteria:

1. Sponsored children are to be from low-income families within walking distance of a project. The official selection criteria state: "When only a percentage of the children are sponsored from an institution, the school or parent committee should choose children among the neediest families for sponsorship."

2. Orphans, children living with a widowed parent or other family member, and refugees are given special priority.

3. The child cannot have been sponsored by another agency.

4. Children from both Christian and non-Christian families may participate equally, but all families must allow their children to participate in the programs Christian religious instruction.

5. Compassion sponsors a maximum of three children per family. Some countries set a lower limit (one or two per family).

6. Children older than 12 years cannot be sponsored. Children in kindergarten and in first, second, or third grade receive top priority; older children (still age 12 or under) receive lower priority.

The last guideline was intended to lengthen the number of years that a child can be sponsored and was fully operational after programs had operated for several years. However, when a program was first introduced into a village, parents tended to select older eligible children rather than younger eligible children. The mean length of sponsorship in our data is 9.3 years.

1.3.3 Survey and Fieldwork

The survey work in the six countries of our main study took place from June 2008 to August 2010. Table 1.3 shows, for each country, a list of villages, rollout years for each village, sample sizes, and survey dates. Some projects started on a large-scale, enrolling up to 100 children in the first year. Others started with fewer children, enrolling only 20 to 30 in the first year. For the larger projects, individuals were randomly selected to be surveyed from the first two or three years of enrollment lists. For smaller projects, data were collected from all children who were enrolled in the first two or three years. To avoid attrition bias, the sample includes both children who were sponsored for many years as well as children who dropped out relatively early.

In some cases the enrollment lists from which we sampled formerly sponsored children were in an electronic database at the country office. In others the only lists were hard copies of computer printouts kept on file at the project sites, which were found after extensive searches of file cabinets or basement boxes. Two of the 19 projects no longer sponsored children through Compassion.

Local assistants were hired to locate the households of formerly sponsored individuals who were on the early enrollment lists. They were usually recommended by project staff, and were known to be responsible, well-respected community members. They also had been raised in the village and so were knowledgeable about the community, but we eschewed hiring enumerators with formal connections with Compassion to avoid bias in responses. We located close to 99% of the targeted households in Uganda, Guatemala, and Kenya, and about 90% in India and Bolivia. In the Philippines slightly less than 80% were located due to a high rate of household mobility following sponsorship. Overall, we located 93.5% of the families of the formerly sponsored children who were on the enrollment lists for the first two to three years the program operated in each village. Families who were not located either had key family members who had passed away, had migrated to unknown areas of the country, or to known areas but without specific details regarding their location.

In addition to these Compassion households, we surveyed 50 to 75 non-Compassion households in each Compassion village, conditional on the presence of an individual in the household being born in the ten years before the Compassion project began operation. We also randomly sampled a similar number of households in neighboring villages without a Compassion program that were similar to the nearby Compassion villages. Households that did not participate in the program, from either Compassion or non-Compassion villages, were surveyed in order to check for intra- household and intra-village spillovers, as explained below in Section 4. The overall, six-country data set includes information on 1,860 formerly sponsored children, 3,704 of their unsponsored siblings, 2,136 individuals from non-participating families in villages where the Compassion program operated, and 2,444 individuals from similar, nearby villages without the Compassion program.

The samples of non-Compassion households, from both the Compassion and the non-Compassion villages, were randomly selected as follows. A starting point in the village was randomly chosen, and then every third household on the street was selected for possible inclusion in the survey. The household was briefly questioned to see whether any of its members met the sponsorship age criteria. When the end of the street or block was reached, the enumerator turned left and continued with every second or third household, then turned right and proceeded in this way, choosing new random points in the village on different days.

Table 1.4 shows summary statistics for the outcome variables, and key control variables. The survey questionnaire we used had questions to collect basic information on adult life outcomes of both sponsored children and their siblings, as well as children in non-participating households. These included questions on an individuals level of schooling, type of employment, and whether he or she held various leadership positions. The questionnaire was designed to ask each question sequentially across all siblings by age to avoid focusing on the sibling(s) who had been sponsored by Compassion. The questions were designed to be easily answered in order to obtain data on the basic life outcomes of adults that would be common knowledge among family members. We eschewed questions that asked for detailed data or for exact values of continuous variables, since family members may not be able to provide such information. Many of the questions were obtained from the education modules in the World Banks Living Standards Measurement Study (LSMS) surveys.

We interviewed all available family members jointly regarding the life outcomes of each formerly sponsored child and his or her siblings. Although, in most cases, several family members answered questions on sibling life outcomes, information was collected on the principal respondent. The most common principal respondents were the parents of the sponsored child (36.6%) and the formerly sponsored child (35.8%), followed by siblings (22.4%) and other relatives (5.2%).

1.4 Empirical Methodology

We employ a variety of estimation techniques to identify the impact of the Compassion sponsorship program. Important to each approach is the programs 12-years-and-under eligibility rule. Figure 1.1 shows the probability that a child in a treated household was sponsored as a function of his or her Age at the Compassion program Introduction (ACI) into his or her village. Averaging over the bars in the figure, a child in a treated household who was between 0 and 12 years old when the program came to his or her village had a 0.458 probability of sponsorship; for a child who was 12-18, the probability was only 0.022. (A few 13-year-old children were sponsored because their photos had been taken and posted for sponsorship when of eligible age, but a sponsor was not found until they were 13.)

Figures 1.2 and 1.3 illustrate differences between treated and non-treated households in years of completed schooling and the probability of completing secondary school, respectively, as a function of two-year ACI categories (for visual smoothness). Both figures compare treated households to non-treated households in Compassion villages. In Figure 1.2, for treated households the difference between the average outcome for those with ACI ≤ 12 and the average for those with ACI ≥ 13 is 1.47 years of schooling. For non-treated households the difference is 0.89, so the difference-in-differences equals 0.58 years. Dividing this figure by the difference in the probability of sponsorship across these two ACI categories (0.436) suggests a 1.33 year impact of being sponsored, not accounting for controls. The analogous difference in differences in Figure 1.3 for secondary school completion is 9.0 percentage points; dividing this by 0.436 suggests that sponsorship increases secondary completion by 20.6 percentage points, which could include spillovers onto other eligible children in Compassion households, but likewise does not adjust for control variables.

To estimate the impact of the program in a manner that more carefully controls for individual and household characteristics, we employ four regression specifications: ordinary least squares (OLS), a generalized method of moments instrumental variables estimator (IV-GMM), and both of these techniques adding household fixed effects (OLS-FE and IV-GMM-FE). These estimation methods must address three concerns to identify the impact of the Compassion program: 1) endogeneity in the selection of households into the program; 2) endogeneity in the selection of children within a particular household into the program; and 3) spillover effects from the program onto non-treated individuals in both Compassion and non-Compassion households in Compassion villages.

To address the first concern, our estimation strategy: (a) allows for an additive, unobserved difference between selected and non-selected households in Compassion villages, and an additive, unobserved difference between households in villages with and without the Compassion program; (b) controls for individual and parental characteristics; and (c) allows for unobserved household fixed effects (for two of the four regression specifications).

The second concern is the possibility of endogenous child selection within families. In practice, Compassion staff often select families for sponsorship, after which the families strongly influence which of their children are sponsored. Endogeneity in the choice of children within a family could bias estimates in either direction; selected children may have been chosen because they seemed of higher ability and so could realize larger gains from the program. Conversely, parents who have preferences for equal outcomes across all their children would select children whom they deem to be more in need of assistance than their siblings. It appeared to us in the field that the latter was more common, which is consistent with Compassions guidelines for field personnel.

We address possible endogeneity in child selection by constructing instrumental variables based on an individuals age and sibling order relative to the year of program rollout in his or her village (the former are ACI categories; the latter are oldest age eligible child, second oldest age eligible child, etc.). These variables should not be related to adult life outcomes except through program participation, and they are highly correlated with the probability of sponsorship; thus they are valid instruments to address endogenous child selection. We observed that needier children tended to be selected for sponsorship from among age-eligible siblings, so we expect any bias in OLS estimates to be downward. Thus *a priori* we expect IV estimates to be larger than OLS estimates.

To address the third concern, spillover effects, each estimation method incorporates a difference-in-differences estimation strategy that allows for measurement of potential spillover effects onto non-sponsored individuals in both Compassion households and non-Compassion households in Compassion villages. Estimation of potential spillover effects rests on two identifying assumptions. The first is that spillovers do not flow from Compassion villages to non-Compassion villages. The second is that spillovers do not trickle from sponsored younger children up to older age-ineligible children in *any* households in Compassion villages.

While there is little reason to expect violations of the first assumption due to the distance between neighboring communities in our study, one objection to the second assumption is the possibility that sponsorship of younger, age-eligible children affects their older, age-ineligible siblings. Yet while there are good reasons to expect positive externalities to trickle down from sponsored children to younger unsponsored siblings, there is also good reason to expect that spillovers onto older, age-ineligible siblings are much smaller, if not zero. Older siblings tend to be less influenced by their younger siblings choices than vice versa. More importantly, education opportunities are usually accessible only within a relevant age range, beyond which older siblings have often passed, preventing them from emulating their younger siblings even if they wanted to do so.

To measure spillover effects onto non-sponsored age-eligible siblings in treated (Compassion) households, we compare differences in life outcomes between age-eligible children and their older age-ineligible siblings in Compassion households with the same differences between these two groups in nearby, non-treated villages. To estimate spillover effects from Compassion onto non-Compassion households in program villages, we compare differences in life outcomes between age-eligible siblings and their older age-ineligible siblings in non-treated households in Compassion villages with the same differences between these two groups in nearby non-Compassion villages. This is done by including dummy variables in the regression analysis that represent each of these groups.

More formally, we assign all individuals who were 16 or younger when the Compassion program began in their villages (or a neighboring village) into seven mutually exclusive categories:

1. Sponsored children, denoted by $T = 1$, who were 12 or younger when the program started in their villages;

2. Siblings of program participants who were 12 or younger when the program was introduced into their villages (denoted by $D_1^{\leq 12} = 1$); they were eligible, but not selected, for the program;

3. Siblings of program participants who were 13-16 when the program was introduced into their villages and thus were ineligible for the program ($D_1^{13-16} = 1$);

4. Individuals in non-Compassion households in program villages who were 12 or younger when the program was introduced $D_2^{\leq 12} = 1$);

5. Individuals in non-Compassion households in program villages who were 13-16 when the program was introduced ($D_2^{13-16} = 1$);

6. Individuals in non-Compassion villages who were 12 or younger when the program was introduced in a neighboring village ($D_3^{\leq 12} = 1$); and finally

7. Individuals in non-Compassion villages who were 13-16 when the program was introduced in a neighboring village ($D_3^{13-16} = 1$).[7]

These categories, and the associated notation, lead to the following regression equation:

$$\begin{aligned} y_i = \alpha_1 D_{1i}^{\leq 12} + \alpha_2 D_{1i}^{13-16} + \tau(D_{1i}^{\leq 12} * T_i) + \beta_1 D_{2i}^{\leq 12} + \beta_2 D_{2i}^{13-16} + \gamma_1 D_{3i}^{\leq 12} + \gamma_2 D_{3i}^{13-16} \\ + \delta C_h + \theta C_v + \mathbf{X_i}\boldsymbol{\phi} + \varepsilon \end{aligned} \tag{1.1}$$

where y_i is the adult outcome of interest for person i, C_h is a dummy variable indicating a household with a sponsored child, C_v is a dummy variable indicating residence in a village with the Compassion program, and $\mathbf{X_i}$ is a vector of controls that include gender, age, age-squared, birth order, number of siblings in a family, and mothers and fathers education.[8]

---

[7]Individuals 17 or older in program households are accounted for by the sum of the program household dummy variable and the program village dummy variable in equation 1.1, and individuals 17 or older in program villages are accounted for by the program village dummy variable. Individuals 17 or older in non-program villages are the omitted category. Note also that $D_1^{\leq 12} = 1$ for both sponsored children and their age-eligible siblings who were not chosen to be sponsored.

[8]A more flexible specification would allow τ to vary for each ACI category (ACI = 12, ACI = 11, etc.) and similarly for α_1, α_2, β_1, β_2, γ_1 and γ_2. We tested the restrictions in equation 1.1 that τ is the same for all sponsored children (i.e. does not vary by ACI) and the analogous restrictions for α_1, α_2, β_1, β_2, γ_1 and γ_2, and the same restrictions in equation 1.4 below, and we cannot reject the hypothesis that these restrictions

This framework allows us to estimate the causal impacts of the Compassion program under different assumptions about spillover effects. If we assume that: i) differences between villages with and without the Compassion program can be fully represented by θC_v and the observed covariates $\mathbf{X}$; ii) differences between participating and non-participating households in villages where Compassion has a program can be represented by δC_h and the observed covariates $\mathbf{X}$; iii) there are no intra-family, intra-village, and inter-village spillover effects; and iv) intra-family child selection is random, then OLS estimates of equation 1.1 would consistently estimate τ, the impact of the sponsorship program on y_i. Moreover, we can test whether these assumptions are reasonable. For example, continuing to assume no spillovers onto individuals with ACI 13, assumptions i) and ii) imply that $\alpha_2 = \beta_2 = \gamma_2$ (C_h, C_v and the observed covariates $\mathbf{X}$ fully account for differences between individuals with ACI of 13-16 in Compassion households, non-Compassion households in Compassion villages and households in non-Compassion villages, so there is no need for these three parameters to differ), and adding assumption iii) implies that $\alpha_1 = \beta_1 = \gamma_1$ (without intra-household or intra-village spillovers, there is also no reason for these three parameters to differ).

Alternatively, if assumptions i) and ii) but assumption iii) does not, intra-household spillovers onto non-treated eligible siblings can be estimated as explained above by the difference-in-differences $[\alpha_1 - \alpha_2] - [\gamma_1 - \gamma_2]$, and intra-village spillovers among children 12 and younger can be estimated by $[\beta_1 - \beta_2] - [\gamma_1 - \gamma_2]$. If spillovers exist, then the full treatment effect is no longer estimated by τ alone. If there are intra-household spillovers, the full program effect on the treated can be estimated by $\tau + [\alpha_1 - \alpha_2] - [\gamma_1 - \gamma_2]$. That is, τ estimates the impact of the program on a treated child relative to his or her siblings of similar age, but since those siblings experienced spillover effects, one needs to add this spillover effect to to obtain the full program impact on the treated child.

---

hold. Unrestricted estimates, as well as the tests of these restrictions, are available from the authors upon request.

Our instrumental variable estimations use a vector of instruments comprised of interactions between dummy variables for a childs age at program introduction (ACI) and dummy variables for sibling order relative to program rollout ($SORR$).[9] These dummy variables have strong predictive power for the child(ren) chosen by parents for the program due to parents tendency to choose the oldest age-eligible siblings for sponsorship.[10] Moreover,they plausibly satisfy the exclusion restriction for equation 1.1, since there is no reason why, after controlling for characteristic variables ($\mathbf{X}$), a childs age at the time of program rollout interacted with sibling order relative to program rollout should affect adult life outcomes except via its effect on the probability of sponsorship.

In addition to aggregating age categories (12 or less and 13-16), we also aggregated our instrumental variables.[11] This is done to avoid potential problems with including large numbers of instruments that individually may be weak, and to provide more reliable asymptotic results given that the sample is divided into 32 clusters (villages). We maintained the distinction between children who were the oldest eligible sibling, the second oldest eligible sibling, and all other eligible siblings, but we grouped individuals by ACI into three categories: 5 or younger, 6-8, and 9-12, yielding nine instruments of the form $\mathbf{Z_{1i}^{ACI}} = (D_{1i}^{ACI} \otimes SORR)$. Other variations in the age aggregation of the instruments yield very similar results.

Thus for IV estimates the first-stage estimation is:

$$\begin{aligned} T_i = \tilde{\alpha}_1 D_{1i}^{\leq 12} + \tilde{\alpha}_2 D_{1i}^{13-16} + \tilde{\beta}_1 D_{2i}^{\leq 12} + \tilde{\beta}_2 D_{2i}^{13-16} + \tilde{\gamma}_1 D_{3i}^{\leq 12} + \tilde{\gamma}_2 D_{3i}^{13-16} \\ + \tilde{\delta} C_h + \tilde{\theta} C_v + \mathbf{X_i}\tilde{\phi} + \mathbf{Z_{1i}^{ACI}}\tilde{\pi} + \varepsilon i \end{aligned} \tag{1.2}$$

---

[9] $SORR$ consists of three dummy variables: oldest sibling among age-eligible siblings at time of program roll-out, second oldest of such siblings at time of program roll-out, and third or higher oldest of such siblings at time of program roll-out.

[10] Among children in Compassion households, the probability of the two oldest age-eligible children being sponsored was 51.5 percent (average over all six countries), compared to only 20.4 percent for all other age-eligible children.

[11] Results are similar when we retain all 54 instrumental variables

and the second stage is:

$$y_i = \alpha_1 D_{1i}^{\leq 12} + \alpha_2 D_{1i}^{13-16} + \tau(D_{1i}^{\leq 12}\hat{T}_i) + \beta_1 D_{2i}^{\leq 12} + \beta_2 D_{2i}^{13-16} + \gamma_1 D_{3i}^{\leq 12} + \gamma_2 D_{3i}^{13-16} + \delta C_h + \theta C_v \boldsymbol{\varepsilon} + \mathbf{X_i}\phi + \varepsilon_i \tag{1.3}$$

where $\hat{T}_{ij}$ is the predicted probability of being sponsored. Equations 1.2 and 1.3 are estimated using generalized method of moments (GMM), which is more efficient than standard IV estimates and allows one to carry out J-tests of over-identifying restrictions to check the validity of the instruments.

We also present OLS and (GMM) IV household-fixed-effect estimates, which control more directly for inter-household unobservables that could affect child selection. The main disadvantage is that the large number of fixed effects may reduce the precision of the estimates. The OLS household fixed-effects (OLS-FE) estimate for child i in household j is:

$$y_{ij} = \alpha_1 D_{1ij}^{\leq 12} + \alpha_2 D_{1ij}^{13-16} + \tau(D_{1ij}^{\leq 12} * T_i) + \beta_1 D_{2ij}^{\leq 12} + \beta_2 D_{2ij}^{13-16} + \gamma_1 D_{3ij}^{\leq 12} + \gamma_2 D_{3ij}^{13-16} + \mathbf{X_{ij}}\boldsymbol{\phi} + \theta_j + \varepsilon_{ij} \tag{1.4}$$

where θ_j is a household fixed effect and other variables are as defined previously. For the IV fixed-effect estimations (IV-GMM-FE), the first stage equations are given by:

$$T_{ij} = \tilde{\alpha}_1 D_{1ij}^{\leq 12} + \tilde{\alpha}_2 D_{1ij}^{13-16} + \tilde{\beta}_1 D_{2ij}^{\leq 12} + \tilde{\beta}_2 D_{2ij}^{13-16} + \tilde{\gamma}_1 D_{3ij}^{\leq 12} + \tilde{\gamma}_2 D_{3ij}^{13-16} + \mathbf{X_{ij}}\tilde{\boldsymbol{\phi}} + \mathbf{Z_{ij}^{ACI}}\tilde{\boldsymbol{\pi}} + \tilde{\theta}_j + \tilde{\varepsilon}_{ij} \tag{1.5}$$

where $\tilde{\theta}_j$ is also a household fixed effect, and $\mathbf{Z_{ij}^{ACI}}$ is the same vector of instruments used in equation 1.4. The second-stage equation of the household fixed-effects estimation is:

$$y_{ij} = \alpha_1 D_{1ij}^{\leq 12} + \alpha_2 D_{1ij}^{13-16} + \tau(D_{1ij}^{\leq 12}\hat{T}_i) + \beta_1 D_{2ij}^{\leq 12} + \beta_2 D_{2ij}^{13-16} + \gamma_1 D_{3ij}^{\leq 12} + \gamma_2 D_{3ij}^{13-16} + \theta_j + \mathbf{X_{ij}}\boldsymbol{\phi} + \varepsilon_{ij} \tag{1.6}$$

1.5 Empirical Results

This section presents estimates of the impact of child sponsorship on completed years of schooling and on the probabilities of completing primary, secondary, and tertiary education. It also summarizes estimates of impacts on employment and leadership, and presents robustness checks.

1.5.1 Estimates of Impact on Education

Table 1.4 provides, by country, descriptive statistics of the outcome variables and the main control variables. Table 1.5 shows how the outcome variables differ between all sponsored children (first column), all non-sponsored children in the sample (middle column), and non-sponsored siblings of sponsored children (last column). Simple t-tests in Table 1.5 that do not account for the influence of control variables show statistically significant differences in all of these adult life outcomes, including formal years of schooling, where sponsored children realize 1.38 more years of schooling than their unsponsored siblings and 1.79 more years of schooling than their unsponsored peers.

Tables 1.6 - 1.9 provide estimates of equations 1.1 through 1.6, that is, estimates based on OLS, IV-GMM, OLS with household-level fixed effects, and IV-GMM with household-level fixed effects. The education outcomes differ for each table; years of completed schooling is in Table 1.6, and primary, secondary, and university completion are in Tables 1.7, 1.8 and 1.9, respectively. OLS estimates of the impact of Compassion sponsorship on completed years of formal schooling in the first column of Table 1.6 show a highly significant estimated direct impact (τ) of a little over one year (1.03 years). The coefficients on the other categories of individuals in the next six rows of the table measure differences in educational outcomes between those individuals and individuals in the same type of household who were 17 or older when the Compassion program was introduced in their village (or a nearby village). For example, the second (third) row suggests that children who were 12 and younger (13-16) in Compassion households when the program was introduced but did not participate in the program eventually attained about 0.42 (0.17) more years of schooling than their siblings who were 17 or older when the program began, but neither difference is statistically significant. Similarly, rows 4 and 5 compare younger household members to those aged 17 and older in non-Compassion households in Compassion villages, and rows 6 and 7 do the same among those in non-Compassion villages. The assumption that the general specification used in equations (1) - (6) is reasonable more specifically that C_h, C_v and the observed covariates in $\mathbf{X}$ adequately account for differences between individuals with ACI of 13-16 across Compassion households, non-Compassion households in Compassion villages and households in non-Compassion villages, which implies that $\alpha_2 = \beta_2 = \gamma_2$ is tested in the eighth row of Table 1.6 and is not rejected.

The results in Table 1.6 also allow one to test for spillover effects, as explained in Section 4. The first check for possible spillovers is to test the assumption of "parallel trends" among non-sponsored students with ACI ≤ 12 in the three types of households, which implies $\alpha_1 = \beta_1 = \gamma_1$; this is not rejected (ninth row). The tenth row of Table 1.6 directly checks for intra-household spillovers by comparing the difference in years of schooling between non-participating individuals with ACI 12 or below and those with ACI 13-16 in Compassion households with the same difference for individuals in nearby non-Compassion villages. The insignificance of this double difference, $(\alpha_1 - \alpha_2) - (\gamma_1 - \gamma_2)$, yields no evidence for intra-household spillovers from sponsored children to age-eligible non-sponsored siblings in Compassion households. The eleventh row tests for spillovers within villages by comparing the difference between individuals with ACI 12 or below and those with ACI 13-16 in non-Compassion households in Compassion villages with the same difference in non-Compassion villages, $(\beta_1 - \beta_2) - (\gamma_1 - \gamma_2)$. Thus there is no evidence of intra-village spillovers from Compassion households onto age-eligible children in non-Compassion households in Compassion villages.

The impact on program participants that includes possible spillover effects, $\tau + (\alpha_1 - \alpha_2) - (\gamma_1 - \gamma_2)$, is in the twelfth row of Table 1.6. This "full" program impact is somewhat smaller than the direct program impact (0.76 vs. 1.03), and is statistically insignificant. Yet the lack of statistical significance is due to adding the four statistically insignificant parameters that attempt to capture spillovers to the estimate of the direct impact. Because the spillover estimates are insignificant, our preferred estimate of the Compassion programs impact on years of schooling is the direct impact (τ), which is 1.03.

The second column of Table 1.6 presents IV estimates of the impact of the Compassion program on years of schooling, which are somewhat higher (direct impact of 1.38 years). Note that the exclusion restrictions are not rejected by the over-identification test (p-value = 0.232), and the F- test of the explanatory power of the excluded instruments is quite large (60.03). However, the Hausman test does not reject the hypothesis that the OLS and IV estimates are equal (p-value = 0.969). As with the OLS estimates, there is no evidence of intra-household or intra-village spillovers, and the "full" program effect on participants (1.34 years) is very similar to the direct effect (1.38), though it is significant only at the 10% level due to the imprecision of the estimated spillover effects.

The third and fourth columns carry out the same estimates in the first two columns, but add household fixed-effects. Estimates of direct impacts (τ) are slightly higher, at 1.12 and 1.46 years, and although the program effect including spillovers, $\tau + (\alpha_1 - \alpha_2) - (\gamma_1 - \gamma_2)$, is estimated to be somewhat higher (1.36 and 1.59 years, respectively), there is no significant evidence of spillovers, so the estimates of the direct impacts in Table 1.6 are our preferred measures of impact.[12]

Estimation results for primary school completion are shown in Table 1.7. The estimates of the direct impact (τ) are generally significant, but small, probably due to the relatively high rate of baseline primary school completion in these villages (88.7%). These estimates are similar across columns, ranging from 4.0 (OLS estimates) to 7.7 percentage points (IV

[12] There is marginally significant evidence that the parallel trends assumption is violated in the IV-GMM-FE specification (p-value of 0.048), but since the seven other tests of parallel trends in Table 6 do not reject that assumption, and the Hausman test does not indicate a need for IV estimation, we do not view this as a major cause for concern.

estimates without fixed effects). All except for the fixed-effect IV estimate in the last column are significant at the 1% level. There is no evidence of spillovers either within households or within Compassion villages.

Table 1.8 presents estimates of the probability of completing secondary school. Because the Compassion program typically sponsors children to the end of secondary school, this is a natural level of school completion to examine. OLS estimates (first column) indicate that the direct effect of the program (τ) raises the probability of completing secondary school by 13.2 percentage points. The IV estimates (second column) show a 16.5 percentage point effect. Neither of these estimates shows significant spillover effects. Household fixed-effects (third and fourth columns), yield marginally significant estimates of household spillovers onto other age-eligible siblings in secondary school completion, about 7 percentage points (p-value 0.109) in the OLS estimates and about 10 percentage points (p-value 0.110) in the IV estimates. (Another indication of spillovers is the rejection of parallel trends for individuals with ACI $\leq$ 12; p-value of 0.011.) Incorporating these estimated spillovers increases the full program effect to 18.5 percentage points (OLS-FE) and 15.9 percentage points (IV- GMM-FE) on sponsored individuals, both of which are highly significant. The spillover point estimates, if valid, also indicate a 7 to 10 percentage point impact on other age-eligible siblings.

Estimated impacts on completion of tertiary education are in Table 1.9. OLS estimates with and without household fixed effects in the first and third columns yield statistically significant direct impacts (τ) of 2.4 and 2.1 percentage points, respectively. Given a baseline completion rate of 4.3%, these point estimates, while small, reflect an approximate 50% increase over this baseline. The IV point estimates in columns 2 and 4 are somewhat larger, but are of low statistical significance; moreover, Hausman tests cannot reject the consistency of the OLS estimates. Point estimates of intra-household spillovers are small, but positive, such that the "full" impacts of the program (including spillovers) range from 3.2 to 5.1 percentage points, each significant at the 5% level. But due to little direct evidence for household spillovers, we consider the more conservative estimates (on τ only) to be the best estimate of program impact. The IV-GMM estimate finds evidence of positive intra-village spillovers that is significant at the 5% level, but since this is the only significant evidence of such spillovers out of 16 estimates in Tables 1.6 - 1.9, it could simply reflect random chance. (At the 5% significance level one would expect 1 out of 20 estimates of a parameter that equals zero to be significantly different from zero.)

Table 1.10 presents estimated direct impacts (τ) separately for each of the six countries, focusing on the OLS and OLS with household fixed effects estimates.[13] While the program impacts are positive and statistically significant in each country, a striking feature of these estimates is the variation across countries. The estimated impacts of sponsorship are highest in Uganda; the OLS estimates are 2.47 years more years of schooling, a 10.6 percentage point increase in primary school completion, and 25.3 and 7.7 percentage point increases for secondary and university completion, respectively (household fixed-effect estimates are quite similar). The second-highest impacts are most often found for Kenya, where point estimates are 1.16 years for formal schooling, with impacts (in percentage points) of 3.3 for primary completion, 12.2 for secondary completion, but no significant impact for university completion. Guatemala and the Philippines have high estimated impacts for secondary school completion, at 14.0 and 11.7 percentage points, respectively. In Bolivia, nearly all of the significant impact occurs at the university level, where sponsored individuals are 5.1 percentage points more likely to acquire a university education. In India, impacts are smaller

[13]We omit country specific instrumental variable estimates for space consideration, and because Hausman tests never reject the consistency of the OLS estimates (with and without household fixed effects) in the combined sample. Also, the GMM-IV results are generally much less precise than the OLS estimates, and this imprecision is worsened by the small samples for each country. Finally GMM IV estimation requires the number of instruments be less than the number of clusters (villages) to obtain estimates of clustered standard errors. This is not an issue for estimates that combine all six countries (which have 32 villages, and use 8 instruments), but for the individual countries we have only 4 to 7 villages.

and confined to the lower levels of education. Overall the magnitude of educational impacts across educational levels is much higher in the African countries than in the Latin American and Asian countries. Thus, the impact of child sponsorship appears to be greatest where counterfactual levels of education are lowest; Kenya and Uganda have the lowest rates of secondary school completion, and are second and third lowest in terms of years of completed schooling, among the six countries in our data.

Table 1.11 shows estimated impacts on total years of formal schooling by country and by gender, based on the OLS fixed-effect specification (OLS estimates without fixed effects are very similar). The most striking feature of these results is that the impacts are generally larger for the gender with the lowest baseline education levels. In particular, in the Philippines and India, where baseline schooling is higher for girls among non-sponsored children, the impact of sponsorship is higher for boys. Similarly, in Uganda, Guatemala, and Bolivia, where baseline schooling is higher for boys among non-sponsored children, point estimates of program impact are higher for girls.

1.5.2 Robustness Checks

We conducted a number of robustness checks on these results. Estimates that limit the sample to those over age 25 yield similar point estimates and significance for the education outcomes. We also tried several different sets of instruments, interacting our instrumental variables with three other variables to obtain (ACI group $\otimes SORR\times$ gender), (ACI group $\otimes SORR\times$ age), (ACI group $\otimes SORR\times$ birth order). Some of these instruments offered higher first-stage F-statistics, others slightly lower. Results from using different instruments yield nearly identical estimates for the educational outcomes. Some instrument sets yielded slightly larger and more significant impacts on employment and leadership outcomes, while others were slightly weaker. Overall, our chosen instrument set is in about the middle of all the instrument sets we tried in terms of both first- and second-stage significance, and thus it provides reasonably stable estimates of the impact of the Compassion program.

A third robustness check verified that the estimated impacts do not reflect negative spillovers onto older siblings. Estimates for the education (and employment) variables on sponsored children who had no older siblings yields coefficients that are generally similar to those provided here. We also considered the possibility of reporting bias that favored the program. To avoid this bias, the enumerators had no formal ties to the sponsorship program. We also find no evidence of reporting bias; separate regressions for each type of principal respondent (parent, sponsored child, sibling, etc.) yield essentially no differences in point estimates or significance of estimated impacts. Lastly, although we found 93.5% of the families of the formerly sponsored children, could the results be significantly different for the 6.5% whom we were unable to locate? This is unlikely since the strongest results are for the two African countries, where we located 99% of the formerly sponsored childrens families.

1.6 Conclusion

We estimate that the Compassion child sponsorship program increases years of completed schooling by 1.03 to 1.46 years over a baseline of 10.19 years, increases the probability of primary school completion by 4.0 to 7.7 percentage points (baseline 88.7%), secondary school completion by 11.6 to 16.5 percentage points (baseline 44.5%), and university completion by 2.1 to 2.4 percentage points (baseline 4.3%).

One can compare these estimated impacts of Compassions child sponsorship program on education outcomes to recent estimates for other educational interventions. For example, Aaronson and Mazumder (2011) examine the introduction of Rosenwald schools from 1913 to 1931 to foster the education among rural blacks in the U.S. south. They conclude that the introduction of these schools increased secondary school completion among rural blacks by 8.3 percentage points and increased formal schooling outcomes by 1.2 years, estimates that are strikingly similar to ours. Our results for the Compassion program compare favorably to the 0.66 year increase in years of schooling found by Schultz (2004) on the PROGRESA/Oportunidades CCT program in Mexico, and the 0.12 to 0.19 increase in years of schooling that Duflo (2001) estimates as the impact of the large school construction program in Indonesia; she also estimates a 6 percentage point increase in primary school completion, an estimate similar to ours, but finds a slightly negative impact on secondary completion.

We conclude by discussing a possible causal mechanism behind these impacts that subsequent work has begun to explore. The most salient characteristic that distinguishes Compassions program from comparable interventions is its emphasis on raising childrens self-esteem, reference points, and aspirations. As such, it aims to simultaneously relieve both internal and external constraints that can impede progress in education.[14] Indeed, the role of psychological factors has gained increasing attention in development economics.[15] Recent research has explored the role of psychological factors on credit decisions (Bertrand et al., 2010), health (Dupas, 2010), technology adoption (Duflo et al., 2011), and education (Kremer et al., 2009).

Three follow-up studies conducted in Bolivia, Kenya, and Indonesia investigate whether adult life outcomes may have been shaped by the Compassion programs focus on developing self-esteem and nurturing aspirations during childhood. Unlike this paper, which examines formerly sponsored subjects who are now adults, these three follow-up studies focus on *currently* sponsored children.

Ross (2010) examines the life aspirations of 270 children living near Compassion sponsorship projects in Bolivia. In response to the question, "What level of education would you say is sufficient for one to be successful today?" sponsored childrens answers yield village-fixed-effect estimates averaging 0.89 years higher than unsponsored children (average $p < 0.05$), a figure just under the estimated impact on years of schooling found in this paper. Spon-

[14]Dalton et al. (2010) provides an excellent theoretical treatment of internalized poverty constraints.

[15]Mullainathan (2006) and Bernard et al. (2011) review the role of psychology in development economics.

sored children also appear to have higher self-expectations for future vocations. When asked "What occupation do you realistically expect to have in the future?", sponsored children were 10.1 to 25.3 percentage points more likely than unsponsored children to respond with a white-collar occupation (average $p < 0.10$).

In a second follow-up study, Ross and Wydick (2011) surveyed 570 children aged 10 to 18 in three villages in Kenya, using an instrumental variable strategy that exploits the same age-eligibility and limited-children-per-family sponsorship rules used in this paper. Estimates indicate that sponsorship raises educational expectations by 0.23 standard deviations of the distribution of those expectations ($p < 0.10$), as well as raising the probability that a child expects to have a white-collar job by 12.5 percentage points. These changes in expectations about future education and vocation are similar to the estimated impacts among the adults found in this paper.

A third follow-up study, Glewwe and Wydick (2012), examines data from 540 poor children in the slums of Jakarta, Indonesia, finding impacts on schooling aspirations from sponsorship that parallel the findings of the previous two studies. In addition, a new box of 24 colored pencils was placed in front of each child, who was then asked to "Draw a picture of yourself in the rain." Using factor analysis with a varimax rotation on twenty dummy variables that relate drawing characteristics to five measures of hopefulness and self-esteem, three latent factors were identifiable: Happiness, Self-Efficacy, and Hopelessness.[16] Regressions controlling for age, gender, family selection, and neighborhood reveal a 0.27 standard deviation increase in happiness among the Compassion- sponsored children ($t = 2.40$), a 0.33 standard deviation increase in self-efficacy ($t = 3.11$), and a 0.52 standard deviation decrease

[16] There is a large literature on interpreting childrens drawings to gauge their psychological well-being (Koppitz, 1968; Klepsch and Logie, 1982; Furth, 2002). Researchers have found empirical correlations between aspects of childs human figure drawings to a variety of professionally diagnosed disorders, including anxiety and emotional insecurity (missing mouth, frowning, use of dark colors), low self-esteem (tiny figure, poor integration of body parts, missing arms or hands), and low self-efficacy (tiny head, short arms).

in hopelessness ($t = 5.19$).

While further work is required to establish a causal link between aspirations and adult life outcomes, the possibility that nurturing aspirations can have important effects on economic development has intriguing implications. Traditionally, development economics–and indeed the practice of economic development–has focused on the relief of external constraints such as school quality, infrastructure, and credit. But it may be that the internal constraints of the poor also contribute to poverty traps in important ways. Further observational and experimental research should seek to better understand the internal constraints faced by the poor, and how development efforts that seek to release internal constraints can complement purely economic interventions and incentives.

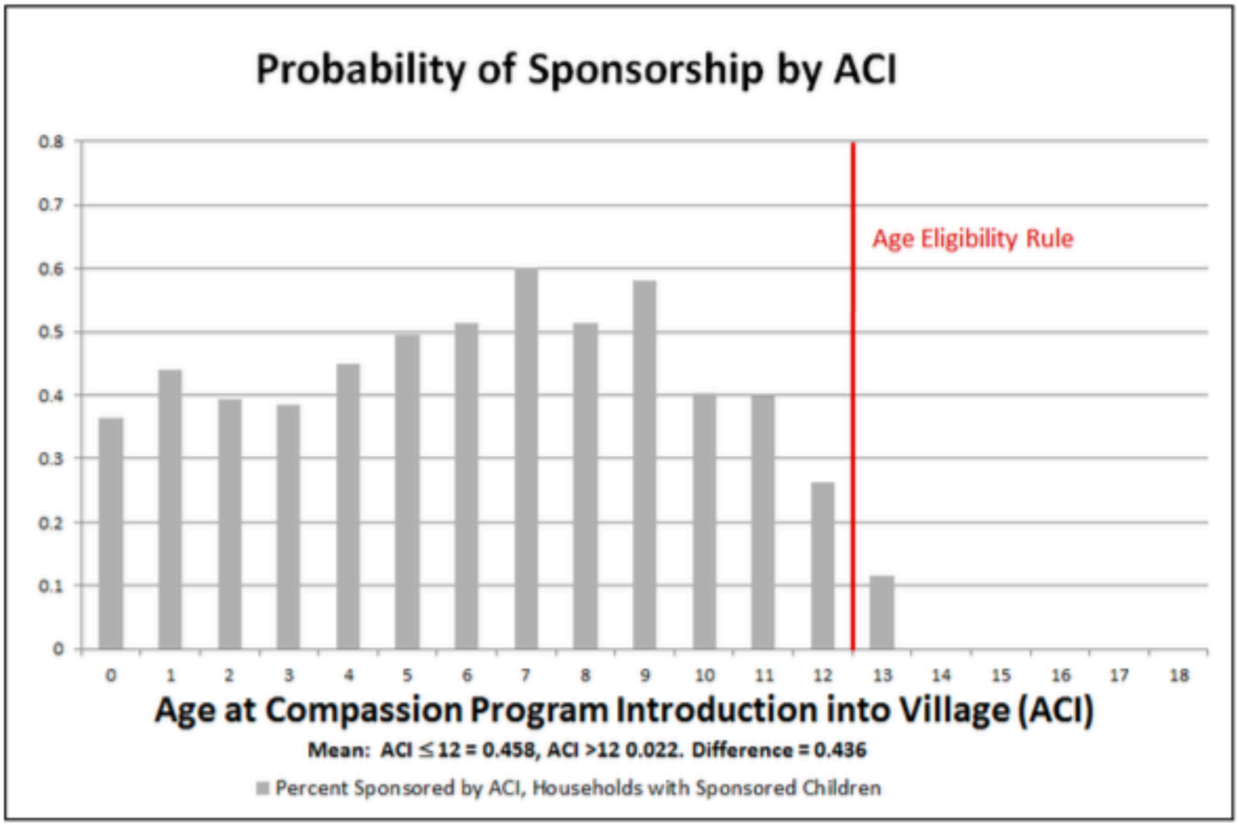

Figure 1.1: Discontinuity in Sponsorship by Age at Time of Program Introduction

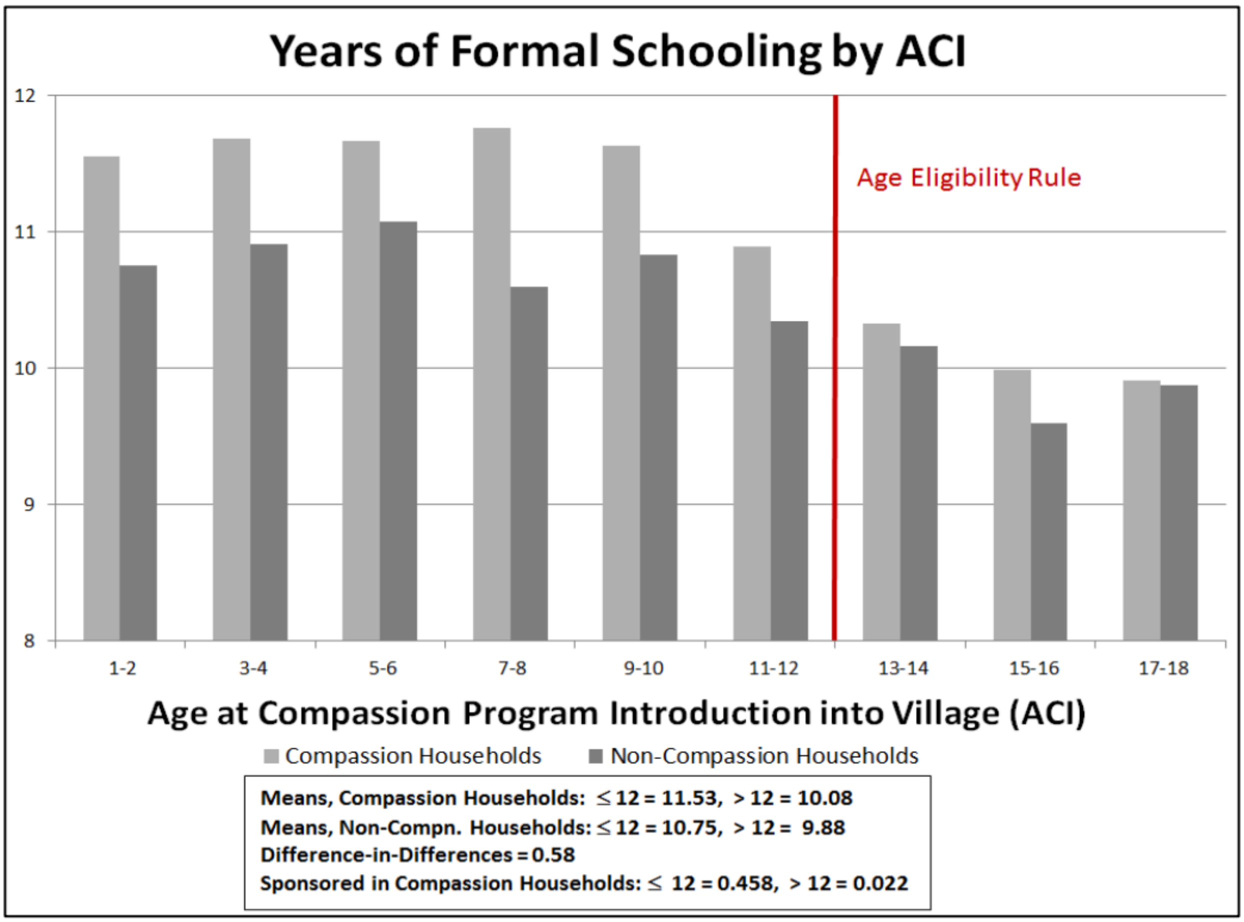

Figure 1.2: Total Years of Formal Schooling by Age when the Compassion Program was introduced into a village (ACI)

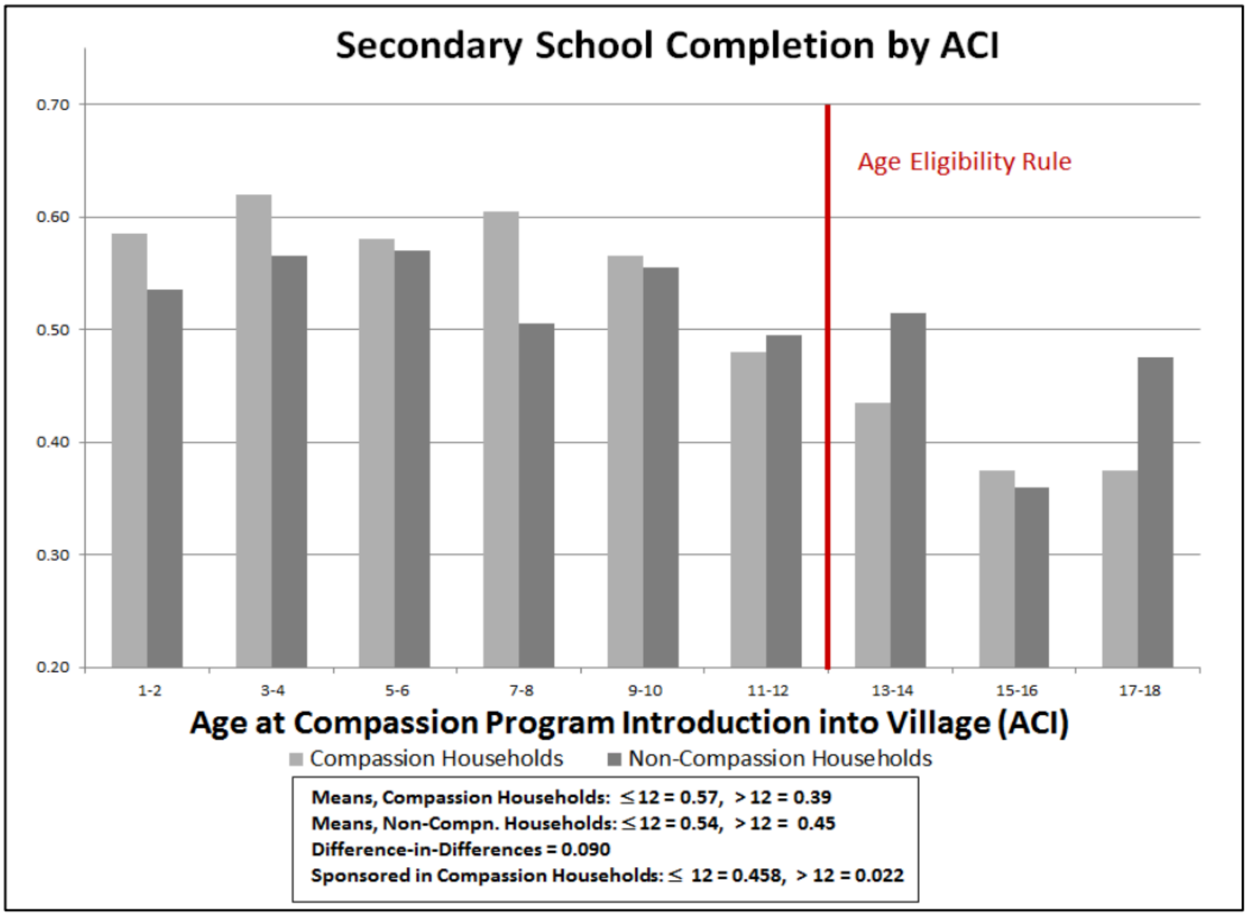

Figure 1.3: Secondary School Completion by Age when the Compassion Program was Introduced into a Village (ACI)

Table 1.1: The Ten Leading International Child Sponsorship Programs

Organization:	**International Headquarters**	**Year Founded**	**No. of Countries**	**Contribution per month**	**No. of Sponsored Children***
1. World Vision†	USA	1953	100	$30	4,100,000
2. Plan USA	USA	1937	49	$24	1,500,000
3. Compassion International†	USA	1952	26	$38	1,288,632
4. ChildFund International	USA	1938	31	$24	510,000
5. Children International	USA	1980	11	$22	340,000
6. Christian Foundation for Children and Aging†	USA	1981	23	$30	291,262
7. Kindernothilfe†	Germany	1959	28	$30	145,814
8. Save the Children	USA	1932	50	$28	120,000
9. SOS Children's Villages	USA	1949	132	$28	80,000
10. Bornefonden	Denmark	1972	5	$34	72,473
Others* (197)					692,979
Total					**9,141,160**

*Child sponsorship organizations by donating country: USA (43), UK (41), France (18), Canada (10), Italy (10), Australia (9), Denmark (7), Spain (7), Norway (6),Germany (5), Sweden (4), Others (16).
† Faith-Based Organization.

Table 1.2: Compassion Program Benefits by Country

Country	**Uniforms**	**Tutoring**	**School Materials**	**Spiritual Instruction**	**Healthcare**	**Gifts from Sponsors**	**Cash to Family**
Uganda	Yes	Yes	Yes	Yes	Yes	Yes	No
Guatemala	Yes	Limited	Yes	Yes	Yes	Yes	No
Philippines	Limited	Limited	Yes	Yes	Yes	Yes	No
India	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Kenya	Yes	Yes	Limited	Yes	Yes	Yes	No
Bolivia	Limited	Yes	Yes	Yes	Yes	Yes	Limited

Table 1.3: Survey Information by Country

Country	**Treatment Villages and Cities (year the program began**	**Non-Treatment Villages and Cities**	**Sample Size**	**Time of Survey**
Uganda	Jinja (1980), Bugiri (1981), Masaka (1989)	Kakooge, Bombo	809	June - August 2008
Guatemala	San Pedro La Laguna (1991), San Juan La Laguna (1992), San Pedro Necta (1992)	San Pablo La Laguna, Santiago Chimaltenango	1,762	May - July 2009
Philippines	Quezon City (1986), Bacolod (1986)	Skybag, Handumanan	1,428	Nov. 2009 - February 2010
India	Tuticorin (1980), Sawyerpuram (1980), Bangalore (1986)	Eral, Bangalore	1,622	March - April 2010
Kenya	Cianthia (1986), Cierria (1986), Nderu (1990), Thigio (1990)	Riakingenyi, Kerwa, Rusigeti	3,056	April - June 2010
Bolivia	Chulla (1992), Los Olivios (1990), Puntiti (1991), Pongonhuyo (1980)	Pairumani-Iscaypata, Igrana	1,467	June - August 2010
Six Countries	**19 Compassion Programs**	**13 Non-Treatment Areas**	**10,144**	**June 2008 - August 2010**

Table 1.4: Summary Statistics

	Uganda	Guatemala	Philippines	India	Kenya	Bolivia	**All Six Countries**
Sponsored as a Child	0.232	0.213	0.168	0.138	0.178	0.198	**0.183**
	(0.423)	(0.409)	(0.374)	(0.345)	(0.383)	(0.398)	**(0.387)**
Years Sponsored	11.325	6.717	7.469	11.065	10.207	9.510	**9.287**
	(3.067)	(2.467)	(4.629)	(3.230)	(3.332)	(3.747)	**(3.790)**
Total Years of Education	9.185	8.859	12.180	11.696	10.442	10.750	**10.566**
	(4.003)	(4.295)	(1.994)	(3.345)	(3.078)	(4.155)	**(3.654)**
Completed Primary	0.841	0.795	0.994	0.937	0.953	0.847	**0.905**
	(0.366)	(0.404)	(0.080)	(0.244)	(0.211)	(0.360)	**(0.293)**
Completed Secondary	0.269	0.486	0.745	0.602	0.338	0.537	**0.485**
	(0.444)	(0.500)	(0.436)	(0.490)	(0.473)	(0.499)	**(0.500)**
Completed Tertiary	0.077	0.022	0.019	0.047	0.040	0.117	**0.049**
	(0.266)	(0.148)	(0.135)	(0.212)	(0.195)	(0.321)	**(0.216)**
Age	28.968	26.635	29.054	32.575	30.806	29.113	**29.827**
	(8.642)	(6.468)	(8.751)	(8.813)	(7.919)	(7.893)	**(8.232)**
Gender	0.476	0.511	0.502	0.500	0.515	0.501	**0.505**
	(0.500)	(0.500)	(0.500)	(0.500)	(0.500)	(0.500)	**(0.500)**
Number of Siblings	4.580	4.838	4.389	3.991	6.307	4.908	**5.076**
	(2.339)	(2.195)	(2.089)	(1.944)	(2.237)	(2.003)	**(2.308)**
Mother's Education	5.385	2.105	9.682	6.982	3.862	2.783	**4.853**
	(3.778)	(2.909)	(3.120)	(3.645)	(4.116)	(3.059)	**(4.338)**
Father's Education	7.024	3.496	10.032	7.629	5.538	4.727	**6.166**
	(4.044)	(3.645)	(3.203)	(3.635)	(4.318)	(3.381)	**(4.310)**

Means; Standard deviations in parentheses

Table 1.5: Summary Means and T-tests of Education Variables

	Mean, Sponsored individuals (std. dev.)	Mean, All non-sponsored individuals (std. dev.)	Difference t-test (std. error)	Mean, Non-sponsored siblings in sponsored households	Difference t-test (std. error)
Total Years of Education	12.03	10.24	1.79***	10.65	1.38***
	(2.790)	(3.74)	(0.093)	(3.41)	(0.092)
Completed Primary	0.984	0.887	0.096***	0.927	0.057***
	(0.127)	(0.316)	(0.008)	(0.260)	(0.006)
Completed Secondary	0.646	0.449	0.196***	0.460	0.185***
	(0.478)	(0.497)	(0.013)	(0.498)	(0.014)
Completed University	0.078	0.043	0.035***	0.049	0.029***
	(0.268)	(0.203)	(0.006)	(0.216)	(0.007)

Full sample size = 10,011. Number of formerly sponsored individuals = 1,834. Number of individuals in non-treated households = 4,560. Number of non-sponsored individuals in treated households = 3,617.

Table 1.6: OLS and IV Estimates of Equations (1), (2), (3), and (4)

	Years of Schooling			
Variables	OLS	IV-GMM	OLS-FE	IV-GMM-FE
Program participant (τ)	1.034***	1.383***	1.118***	1.455***
	(0.152)	(0.441)	(0.121)	(0.407)
Compassion hh: ACI $\leq$ 12 (α_1)	0.421	-0.069	0.720	0.765
	(0.555)	(0.330)	(0.432)	(0.475)
Compassion hh: ACI 13-16 (α_2)	0.169	-0.236	0.039	0.082
	(0.417)	(0.251)	(0.265)	(0.207)
Program village ACI $\leq$ 12 (β_1)	0.588	0.476	0.897*	0.916
	(0.611)	(0.437)	(0.436)	(0.579)
Program village ACI 13-16 (β_2)	-0.013	-0.205	0.414	0.405
	(0.533)	(0.430)	(0.459)	(0.468)
Nonprog. village ACI $\leq$ 12 (γ_1)	0.690	0.248	-0.085	0.047
	(0.734)	(0.697)	(0.504)	(0.435)
Nonprog. village ACI 13-16 (γ_2)	0.161	0.041	-0.525**	-0.499***
	(0.523)	(0.486)	(0.197)	(0.185)
F-tests/χ^2 tests of parallel trends:				
$\alpha_2 = \beta_2 = \gamma_2$	0.08	0.26	2.26	6.08**
	[0.926]	[0.877]	[0.121]	[0.048]
$\alpha_1 = \beta_1 = \gamma_1$	0.07	1.54	1.03	1.38
	[0.935]	[0.464]	[0.368]	[0.501]
Intra-Household Spillovers				
$(\alpha_1 - \alpha_2) - (\gamma_1 - \gamma_2)$	-0.277	-0.039	0.240	0.137
	(0.686)	(0.686)	(0.516)	(0.555)
Intra-Village Spillovers				
$(\beta_1 - \beta_2) - (\gamma_1 - \gamma_2)$	0.072	0.474	0.043	-0.036
	(0.631)	(0.584)	(0.540)	(0.529)
Program Impact				
including household spillovers				
$\tau + (\alpha_1 - \alpha_2) - (\gamma_1 - \gamma_2)$	0.757	1.344*	1.359**	1.592***
	(0.678)	(0.713)	(0.506)	(0.446)
Hausman Test [p-value]		[0.969]		[0.922]
Over-identification (J-test)		9.30		6.05
[p-value]		[0.232]		[0.534]
Weak IV Test				
(F-statistic)		60.03		48.93
Observations	9954	9954	9954	9954
R-squared	0.253	0.250	0.064	0.047

***Significant at 1%; **significant at 5%; *significant at 10%. Clustered standard errors at village level in parentheses. Each regression includes controls for age at program introduction (ACI) at age 12 and below and age 13-16, age, age-squared, birth order, gender, status as oldest child, mother's education, father's education, mother's education missing and father's education missing.

Table 1.7: OLS and IV Estimates of Equations (1), (2), (3), and (4)

		Primary School Completion		
Variables	OLS	IV-GMM	OLS-FE	IV-GMM-FE
Program participant (τ)	0.040***	0.077***	0.049***	0.050
	(0.09)	(0.028)	(0.010)	(0.045)
Compassion hh: ACI $\leq$ 12 (α_1)	0.017	-0.013	0.020	0.034
	(0.054)	(0.029)	(0.038)	(0.028)
Compassion hh: ACI 13-16 (α_2)	0.004	-0.024	-0.010	0.007
	(0.044)	(0.030)	(0.026)	(0.018)
Program village ACI $\leq$ 12 (β_1)	0.037	0.006	0.066**	0.047
	(0.057)	(0.049)	(0.031)	(0.038)
Program village ACI 13-16 (β_2)	-0.0217	-0.051	0.021	0.012
	(0.048)	(0.040)	(0.031)	(0.034)
Nonprog. village ACI $\leq$ 12 (γ_1)	0.066	-0.008	0.013	-0.006
	(0.077)	(0.071)	(0.049)	(0.045)
Nonprog. village ACI 13-16 (γ_2)	0.002	0.016	-0.029	-0.036*
	(0.055)	(0.048)	(0.021)	(0.019)
F-tests/χ^2 tests of parallel trends:				
$\alpha_2 = \beta_2 = \gamma_2$	0.16	1.20	0.98	3.20
	[0.856]	[0.548]	[0.388]	[0.202]
$\alpha_1 = \beta_1 = \gamma_1$	0.17	0.22	0.74	0.72
	[0.8465]	[0.894]	[0.487]	[0.698]
Intra-Household Spillovers				
$(\alpha_1 - \alpha_2) - (\gamma_1 - \gamma_2)$	-0.051	0.035	-0.016	-0.001
	(0.072)	(0.071)	(0.053)	(0.053)
Intra-Village Spillovers				
$(\beta_1 - \beta_2) - (\gamma_1 - \gamma_2)$	-0.006	0.081	0.003	0.005
	(0.069)	(0.067)	(0.057)	(0.045)
Program Impact				
including household spillovers				
$\tau + (\alpha_1 - \alpha_2) - (\gamma_1 - \gamma_2)$	-0.011	0.112	0.033	0.050
	(0.073)	(0.075)	(0.053)	(0.043)
Hausman Test [p-value]		[0.888]		[0.830]
Over-identification (J-test)		9.47		8.23
[p-value]		[0.220]		[0.313]
Weak IV Test				
(F-statistic)		60.03		48.93
Observations	9954	9954	9954	9954
R-squared	0.123	0.104	0.031	0.030

***Significant at 1%; **significant at 5%; *significant at 10%. Clustered standard errors at village level in parentheses. Each regression includes controls for age at program introduction (ACI) at age 12 and below and age 13-16, age, age-squared, birth order, gender, status as oldest child, mother's education, father's education, mother's education missing and father's education missing.

Table 1.8: OLS and IV Estimates of Equations (1), (2), (3), and (4)

		Secondary School Completion		
Variables	OLS	IV-GMM	OLS-FE	IV-GMM-FE
Program participant (τ)	0.132***	0.165*	0.116***	0.062
	(0.028)	(0.094)	(0.017)	(0.085)
Compassion hh: ACI $\leq$ 12 (α_1)	-0.027	-0.057	0.093*	0.167***
	(0.053)	(0.058)	(0.055)	(0.064)
Compassion hh: ACI 13-16 (α_2)	-0.024	-0.050*	-0.001	0.022
	(0.039)	(0.027)	(0.034)	(0.024)
Program village ACI $\leq$ 12 (β_1)	-0.011	-0.009	0.113	0.166
	(0.059)	(0.039)	(0.075)	(0.109)
Program village ACI 13-16 (β_2)	-0.032	-0.055	0.055	0.081
	(0.062)	(0.045)	(0.076)	(0.087)
Nonprog. village ACI $\leq$ 12 (γ_1)	-0.022	-0.013	-0.032	-0.008
	(0.060)	(0.058)	(0.050)	(0.045)
Nonprog. village ACI 13-16 (γ_2)	-0.014	-0.008	-0.062*	-0.056*
	(0.045)	(0.043)	(0.031)	(0.029)
F-tests/χ^2 tests of parallel trends:				
$\alpha_2 = \beta_2 = \gamma_2$	0.03	0.83	1.37	6.56**
	[0.970]	[0.660]	[0.270]	[0.038]
$\alpha_1 = \beta_1 = \gamma_1$	0.04	0.53	2.33	8.95**
	[0.962]	[0.765]	[0.114]	[0.011]
Intra-Household Spillovers				
$(\alpha_1 - \alpha_2) - (\gamma_1 - \gamma_2)$	0.005	-0.003	0.069*	0.097
	(0.055)	(0.065)	(0.421)	(0.060)
Intra-Village Spillovers				
$(\beta_1 - \beta_2) - (\gamma_1 - \gamma_2)$	0.029	0.051	0.027	0.036
	(0.056)	(0.048)	(0.048)	(0.053)
Program Impact				
including household spillovers				
$\tau + (\alpha_1 - \alpha_2) - (\gamma_1 - \gamma_2)$	0.137**	0.162**	0.185***	0.159***
	(0.058)	(0.064)	(0.043)	(0.050)
Hausman Test [p-value]		[0.959]		[0.785]
Over-identification (J-test)		6.18		6.85
[p-value]		[0.519]		[0.444]
Weak IV Test				
(F-statistic)		60.03		48.93
Observations	9954	9954	9954	9954
R-squared	0.164	0.161	0.033	0.019

***Significant at 1%; **significant at 5%; *significant at 10%. Clustered standard errors at village level in parentheses. Each regression includes controls for age at program introduction (ACI) at age 12 and below and age 13-16, age, age-squared, birth order, gender, status as oldest child, mother's education, father's education, mother's education missing and father's education missing.

Table 1.9: OLS and IV Estimates of Equations (1), (2), (3), and (4)

	University Completion			
Variables	OLS	IV-GMM	OLS-FE	IV-GMM-FE
Program participant (τ)	0.024***	0.036*	0.021**	0.050
	(0.009)	(0.021)	(0.009)	(0.035)
Compassion hh: ACI $\leq$ 12 (α_1)	0.016	0.009	0.033*	0.022
	(0.014)	(0.011)	(0.017)	(0.016)
Compassion hh: ACI 13-16 (α_2)	0.009	0.004	0.015	0.014
	(0.012)	(0.010)	(0.013)	(0.012)
Program village ACI $\leq$ 12 (β_1)	0.030*	0.030*	0.024	0.015
	(0.017)	(0.012)	(0.023)	(0.028)
Program village ACI 13-16 (β_2)	0.011	0.006	0.013	0.004
	(0.019)	(0.012)	(0.018)	(0.019)
Nonprog. village ACI $\leq$ 12 (γ_1)	-0.004	-0.001	0.002	0.001
	(0.016)	(0.014)	(0.020)	(0.019)
Nonprog. village ACI 13-16 (γ_2)	-0.003	0.000	-0.007	-0.007
	(0.016)	(0.015)	(0.011)	(0.011)
F-tests/χ^2 tests of parallel trends:				
$\alpha_2 = \beta_2 = \gamma_2$	0.24	0.11	1.17	1.99
	[0.786]	[0.948]	[0.324]	[0.370]
$\alpha_1 = \beta_1 = \gamma_1$	1.60	8.23**	1.07	0.96
	[0.218]	[0.016]	[0.356]	[0.619]
Intra-Household Spillovers				
$(\alpha_1 - \alpha_2) - (\gamma_1 - \gamma_2)$	0.008	0.006	0.009	0.002
	(0.013)	(0.013)	(0.013)	(0.016)
Intra-Village Spillovers				
$(\beta_1 - \beta_2) - (\gamma_1 - \gamma_2)$	0.020	0.025**	0.002	0.004
	(0.015)	(0.012)	(0.021)	(0.021)
Program Impact				
including household spillovers				
$\tau + (\alpha_1 - \alpha_2) - (\gamma_1 - \gamma_2)$	0.032**	0.041**	0.030**	0.051**
	(0.014)	(0.018)	(0.015)	(0.025)
Hausman Test [p-value]		[0.625]		[0.304]
Over-identification (J-test)		9.45		2.73
[p-value]	[0.222]		[0.909]	
Weak IV Test				
(F-statistic)		60.03		48.93
Observations	9954	9954	9954	9954
R-squared	0.021	0.020	0.012	0.006

***Significant at 1%; **significant at 5%; *significant at 10%. Clustered standard errors at village level in parentheses. Each regression includes controls for age at program introduction (ACI) at age 12 and below and age 13-16, age, age-squared, birth order, gender, status as oldest child, mother's education, father's education, mother's education missing and father's education missing.

Table 1.10: OLS and OLS Household Fixed Effect Estimates for Education by Country

Variables						
TOTAL YEARS OF EDUCATION	**Uganda**	**Guatemala**	**Philippines**	**India**	**Kenya**	**Bolivia**
Program participant (τ), OLS	2.472***	0.528***	0.573***	0.658**	1.156***	0.668**
	(0.236)	(0.113)	(0.064)	(0.226)	(0.142)	(0.184)
R^2	0.332	0.426	0.213	0.246	0.156	0.307
Program participant (τ), HH FE	2.216***	0.830***	0.631***	0.768**	1.312***	0.733**
	(0.232)	(0.166)	(0.008)	(0.171)	(0.141)	(0.228)
R^2	0.127	0.147	0.114	0.062	0.075	0.128
Baseline, Untreated	*8.37*	*8.12*	*12.11*	*11.45*	*10.22*	*10.32*
PRIMARY COMPLETION	**Uganda**	**Guatemala**	**Philippines**	**India**	**Kenya**	**Bolivia**
Program participant (τ), OLS	0.106**	-0.011*	0.003	0.030*	0.033*	0.380
	(0.028)	(0.005)	(0.002)	(0.012)	(0.015)	(0.030)
R^2	0.206	0.329	0.060	0.093	0.036	0.231
Program participant (τ), HH FE	0.097**	0.027*	0.005	0.029*	0.052**	0.049
	(0.034)	(0.012)	(0.003)	(0.013)	(0.016)	(0.031)
R^2	0.076	0.120	0.009	0.027	0.014	0.170
Baseline, Untreated	*0.795*	*0.744*	*0.993*	*0.926*	*0.948*	*0.816*
SECONDARY COMPLETION	**Uganda**	**Guatemala**	**Philippines**	**India**	**Kenya**	**Bolivia**
Program participant (τ), OLS	0.253**	0.140***	0.117***	0.083*	0.122***	0.013
	(0.057)	(0.011)	(0.018)	(0.046)	(0.021)	(0.014)
R^2	0.196	0.293	0.151	0.195	0.095	0.204
Program participant (τ), HH FE	0.218**	0.132***	0.120***	0.087*	0.115**	0.026
	(0.058)	(0.021)	(0.003)	(0.046)	(0.035)	(0.024)
R^2	0.075	0.099	0.060	0.039	0.037	0.052
Baseline, Untreated	*0.192*	*0.406*	*0.732*	*0.568*	*0.314*	*0.503*
UNIVERSITY COMPLETION	**Uganda**	**Guatemala**	**Philippines**	**India**	**Kenya**	**Bolivia**
Program participant (τ), OLS	0.077**	0.008	0.002	0.001	0.011	0.051*
	(0.019)	(0.021)	(0.009)	(0.029)	(0.011)	(0.025)
R^2	0.107	0.063	0.024	0.050	0.058	0.083
Program participant (τ), HH FE	0.078***	0.021	0.007	-0.002	0.006	0.053*
	(0.015)	(0.011)	(0.009)	(0.024)	(0.005)	(0.022)
R^2	0.037	0.030	0.045	0.015	0.015	0.036
Baseline, Untreated	*0.055*	*0.016*	*0.020*	*0.042*	*0.037*	*0.102*
Observations	**809**	**1656**	**1390**	**1591**	**3050**	**1458**

***Significant at 1%; **significant at 5%; *significant at 10%. Clustered standard errors at village level in parentheses. Each regression includes controls for age at program introduction (ACI) at age 12 and below and age 13-16, age, age-squared, birth order, gender, status as oldest child, mother's education, father's education, mother's education missing, and father's education missing.

Table 1.11: OLS Fixed-Effect Estimates for Total Years of Schooling by Country and Gender

Variables	BOYS					
TOTAL YEARS OF EDUCATION	**Uganda**	**Guatemala**	**Philippines**	**India**	**Kenya**	**Bolivia**
Program participant (τ)	1.463**	0.496	0.706***	0.671**	1.486***	0.481***
	(0.518)	(0.361)	(0.041)	(0.229)	(0.189)	(0.104)
Observations	385	848	697	795	1569	730
R^2	0.169	0.141	0.071	0.064	0.069	0.134
Baseline, Untreated	*8.39*	*8.48*	*11.86*	*11.24*	*10.29*	*10.55*
	GIRLS					
TOTAL YEARS OF EDUCATION	**Uganda**	**Guatemala**	**Philippines**	**India**	**Kenya**	**Bolivia**
Program participant (τ)	2.735***	1.190***	0.395***	0.423	0.935***	0.694
	(0.405)	(0.173)	(0.053)	(0.328)	(0.124)	(0.517)
Observations	424	808	693	796	1481	728
R^2	0.163	0.214	0.154	0.110	0.081	0.165
Baseline, Untreated	*8.35*	*7.74*	*12.37*	*11.66*	*10.14*	*10.08*

***Significant at 1%; **significant at 5%; *significant at 10%. Estimations include fixed-effects at the household level. Clustered standard errors at the village level in parentheses. Each regression includes controls for age at program introduction (ACI) at age 12 and below and age 13-16, age, age-squared, birth order, gender, and status as oldest child.

Chapter 2

DOES CHILD SPONSORSHIP PAY OFF IN ADULTHOOD? AN INTERNATIONAL STUDY OF IMPACTS ON INCOME AND WEALTH

The published version of this paper can be found in *World Bank Economic Review* (2017) 31 (2): 434-458. DOI: https://doi.org/10.1093/wber/lhv081. University of Oxford Press.

Bruce Wydick, Paul Glewwe, and Laine Rutledge

Abstract

We estimate the impact of international child sponsorship on adult income and wealth of formerly sponsored children using data on 10,144 individuals in six countries. To identify causal effects, we utilize an age-eligibility rule followed from 1980 to 1992 that limited sponsorship to children 12 years old or younger when the program was introduced in a village, allowing comparisons of sponsored children with older siblings who were slightly too old to be sponsored. Estimations indicate that international child sponsorship raised monthly income by approximately $15 over an untreated baseline of $75, principally from increasing future labor market participation. Formerly sponsored boys become more likely to have careers as teachers, tech workers, and hold salaried blue collar jobs as adults. Formerly sponsored girls are also more likely to enter the labor market, particularly as clerical workers, nurses, tech workers, and in careers in finance and business. We find evidence for positive impacts on overall dwelling quality in adulthood, sponsorship causing a 3 percentage point increase in the probability of living in a home with electricity. Sponsorship also significantly increases the probability of living in a rented or owned home independently from parents as an adult. Our results show modest evidence of impacts on adult ownership of consumer durables,

mainly limited to increased ownership of mobile phones. Finally, we find modest effects of child sponsorship on reduced childbearing in adulthood, but insignificant impacts on age of marriage.

2.1 Introduction

Millions of households in wealthy countries support non-profit organizations whose aim is to alleviate poverty in the developing world. But only recently has a growing body of research in development economics begun to rigorously evaluate the impact of these programs on their intended beneficiaries.[1] International child sponsorship is one of the most popular approaches taken by ordinary households in wealthy countries to help impoverished children overseas. We estimate that there are currently 9.14 million sponsored children in the world today, the vast majority of whom are sponsored by individuals and families in wealthy countries.[2] Donors typically contribute $25-40 per month to sponsor a child. In many cases child sponsorship organizations use these funds to provide school uniforms, tuition, nutritious meals, and programming that directly benefits sponsored children. Other types of sponsorship programs pool funds to invest in programming and infrastructure that benefits children in the community more broadly.[3]

For many individuals, child sponsorship represents their most direct contact with the

---

[1] See for example Cristia et al. (2012) evaluating the One Laptop Per Child Program, Rawlins et al. (2014) on the nutritional impacts of dairy cows and meat goats donated via the Heifer Project, and the analysis of Blattman et al. (2014) on cash transfers.

[2] We estimate this figure based on comprehensive internet search across multiple languages for sponsorship programs. For details on how the 9.12 million figure was compiled, see Wydick, Glewwe and Rutledge (2013b).

[3] Of the top ten child sponsorship organizations, a more direct child-centered approach is taken by Compassion International, ChildFund, Children International, CFCA, and Bornefonden. The community-centered approach is favored by World Vision, Plan International, Kindernothilfe, Save the Children, and SOS Children's Villages.

poor in developing countries. Donors are drawn to child sponsorship because of the personalization of the relationship between sponsor and child. But whether child sponsorship actually benefits sponsored children has remained an open question. In Wydick, Glewwe and Rutledge (2013b), a companion paper to this research, we find international child sponsorship to have a statistically significant and positive effect on educational outcomes in all six survey countries (Bolivia, Guatemala, Kenya, Uganda, India, and the Philippines). Sponsorship during childhood increased the probability of secondary school completion by 12-18 percentage points over a 44.5% baseline, and increased completed years of schooling by 1.03-1.45 years. Sponsorship also increased adult white-collar employment by 6.5 percentage points over an 18.5% baseline as well as the probability of being a community leader.

Previous research has studied the impacts of various programs on children's persistence in school in developing countries. Examples include Drèze and Kingdon (2001) and Kremer and Vermeersch (2004) who find positive impacts of school meal programs on school attendance in India and Kenya, respectively. In a randomized trial, Evans, Kremer and Ngatia (2008) find a nearly 40% reduction in absenteeism from the random provision of free school uniforms, while Kremer, Miguel and Thornton (2009) estimate that a merit scholarship program for girls boosted attendance by 5 percentage points. Aside from our research, the only other investigation related to ascertaining the impacts of child sponsorship is Kremer, Moulin and Namunyu (2003). In this paper the authors assess the impact of a Dutch child-sponsorship program, finding that even a relatively low-cost program focused on the provision of school uniforms and textbooks to each child caused sponsored children to advance a third of a grade farther in schooling completion.

In this paper we present results for the impacts of child sponsorship on the adult income[4]

[4]Our study examines changes in labor income, which includes wages paid by an employer, income earned by an entrepreneur from a small enterprise, or income from farming. We do not study income earned from capital holdings, as these were deemed to be insignificant for the great majority of the households in our sample. In our study, the terms "labor income" and "wages" have a similar meaning but are not exactly the

and wealth of children sponsored through one of the leading international child sponsorship organizations. An understanding of these impacts is important for the millions of individuals in wealthy countries involved in international child sponsorship, individuals who are likely to view their contributions as an investment in these overseas children that yields tangible economic returns in the future, when they are adults. But it is also important for governments in countries implementing similar programs that work directly with impoverished children, helping them to understand whether direct investments in child development are financially sustainable by virtue of the positive impacts on the future incomes of beneficiaries. Thus we ask the question: Does international child sponsorship pay off for children in adulthood?

2.2 Methodology

Fieldwork for our six-country study took place from 2008 to 2010. We obtained initial enrollment lists from village projects that were rolled out from 1980 to 1992 by Compassion International, the world's third largest sponsorship organization, which currently sponsors approximately 1.4 million children in 26 countries. Compassion's child sponsorship program is a very intensive intervention in the lives of impoverished children. Children typically begin sponsorship at age 4-6 and continue into their mid-teens. Many sponsored children attend retreats together with program staff that focus on the nurture of their spiritual and moral values as well as their aspirations. Since even in a typical week children typically spend about 8-10 hours per week after school and on Saturdays participating in the program, and because the average duration of sponsorship is 9.3 years, this means that during the course of their childhood, on average sponsored children spend slightly more than 4,000 hours participating in Compassion programming. Average years of participation by country is given in Table

---

same; we use the term "wages" as conditional on working status and "labor income" in contexts that are unconditional on working status, *e.g.* labor income increases when one enters the workforce and begins to earn a wage.

2.1.

Across the countries in which it is implemented, the Compassion program contains many similar elements. In each country, Compassion uses funding to provide tuition fees for children, several nutritious meals per week, basic healthcare, school uniforms, and an after-school tutoring program. The tutoring program not only helps sponsored children with homework and gives them additional academic instruction, but emphasizes spiritual and character formation and the development of schooling and vocational aspirations and self-esteem. Note that vocational training was not included in any of these six countries.

The program has changed slightly since the time of the study and varies somewhat between countries. In the past, Compassion worked in tandem with local schools (which is true of our data from Guatemala and the Philippines) but more recently has operated through local church-affiliated tutoring centers. We study the impacts from children involved in the program in India and Bolivia when small cash transfers were given to parents of participating children, but not in other countries. Other than these differences the program across countries is highly standardized.

Through the use of local enumerators, we were able to locate 93.5% of the families of these formerly sponsored children, who by the time of the survey were aged 17 to 43. Our field personnel were unaffiliated with Compassion, in order to reduce bias in the responses of our subjects. We administered our survey first-hand to households of formerly sponsored children, a random sample of non-participating households in 19 program villages, and a random sample of households in 13 neighboring, non-program villages. The survey questionnaire was administered to family members (typically parents or adult siblings) present at the time of the survey, and data were obtained on all grown siblings in the household cohort, including the non-sponsored siblings of sponsored children. We also administered the survey to 50-75 randomly selected households with children in a similar cohort age that did not participate in the program in program villages, as well as 50-75 randomly selected households with children in a similar cohort age in nearby non-program villages.

Overall, our data contain information on educational and vocational outcomes, monthly labor income, consumer good ownership, and dwelling quality on 1,860 formerly sponsored children, 3,704 of their unsponsored siblings, 2,136 individuals of a similar age from non-participating families in villages where the Compassion program operated, and 2,444 individuals from similar, nearby villages without the Compassion program.

There are several empirical challenges to estimating the program's causal effects on future income and wealth. First, there may be non-random selection of households with eligible children into the program. Second, since a limited number of children per household were eligible for sponsorship (ranging from one in the African countries to three in the Latin American countries), intra-household selection of children for sponsorship may not be random. Third, there may be spillover effects from sponsored children onto their siblings, or onto other children in the village, which complicates the estimation. Lastly, when estimating impacts on future labor income, impacts on employment must be separated from impacts on wages, conditional on employment.

To identify causal effects of child sponsorship, we use a program age-eligibility requirement, which stipulated that only children 12 and under could enter the program in any year, including the program's first year in a village. Figure 2.1 shows the program's strong adherence to this rule. Because a child's age at the time of program rollout in his or her village is independent of adult life outcomes, except via its impact on program participation, we can use the age-eligibility rule as an instrumental variable that allows one to account (and test) for non-random intra-household selection of children for sponsorship. To address possible endogenous household selection into the program, we present estimates with household fixed effects, which control for unobserved differences in parenting behavior and household environments.[5] Implicitly, this compares life outcomes of children who were age-eligible for

[5] We found that program staff usually selected households for participation, and then parents chose which

sponsorship with their siblings who were too old for sponsorship when the program arrived in their village.

The regression estimates allow for the possibility of spillovers. Dummy variables are included for: (a) Sponsored children, who were 12 or younger when the program started in their villages (denoted by $T = 1$); (b) Program participants' siblings who were 12 or younger when the program began in their villages and, while eligible, were not selected for sponsorship (denoted $D_1^{<=12} = 1$); (c) Program participants' siblings age 13-16 when the program arrived in their villages, and thus were ineligible ($D_1^{13-16} = 1$); (d) Individuals in non-Compassion households in program villages who were 12 or younger at program introduction ($D_2^{<=12} = 1$); (e) Individuals in non-Compassion households in program villages age 13-16 at program introduction ($D_2^{13-16} = 1$); (f) Individuals 12 or younger in non-Compassion villages when the program started in a neighboring village ($D_3^{<=12} = 1$); and (g) Individuals 13-16 in non-Compassion villages when the program began in a neighboring village ($D_3^{13-16} = 1$). Individuals 17 or older in non-program villages are the omitted category.

The household fixed-effects equation for child i in household j is

$$\begin{aligned} y_{ij} = \alpha_1 D_{1ij}^{<=12} + \alpha_2 D_{1ij}^{13-16} + \tau(D_{1ij}^{<=12} * T_i) + \beta_1 D_{2ij}^{<=12} + \beta_2 D_{2ij}^{13-16} \\ + \gamma_1 D_{3ij}^{<=12} + \gamma_2 D_{3ij}^{13-16} + \mathbf{X_{ij}}\varphi + \theta_j + \varepsilon_{ij} \end{aligned} \tag{2.1}$$

where y_{ij} measures labor income or wealth, $\mathbf{X_{ij}}$ is a vector of controls that include age, gender, birth order and oldest child, and θ_j is a household fixed effect. Assuming that spillovers: (a) occur only within villages; and (b) affect age-eligible, but not age-ineligible, siblings of sponsored children, then $[\alpha_1 - \alpha_2] - [\gamma_1 - \gamma_2]$ captures spillovers from sponsored children onto non-sponsored siblings age 12 and younger and $[\beta_1 - \beta_2] - [\gamma_1 - \gamma_2]$ captures spillovers onto age-eligible children in non-Compassion households in program villages.[6]

children to be sponsored.

[6]Spillovers onto non-sponsored siblings may reflect extra income available from sponsorship, role model effects, and parental reallocation of assistance to non-sponsored children.

To address possibly endogenous selection of age-eligible children within families we use instrumental variable estimation. The oldest age-eligible sibling was sponsored most often, followed by the second-oldest age-eligible sibling, third oldest, etc., so the instruments are interaction terms between three age-at-program-rollout categories (4 years and under, 5-8, 9-12) and dummy variables for oldest age-eligible sibling, second-oldest age-eligible sibling, and younger age-eligible siblings, yielding a vector of nine instruments.[7] In the first stage, the probability of sponsorship, $\hat{T}_i$, is estimated using these instruments and the vector of controls; $\hat{T}_i$ replaces the treatment variable (T_i) in (2.1) in a second-stage regression.

Estimating the impact of sponsorship on monthly wages involves another challenge: wages are unobserved for the 61% in the sample who were not working. This suggests the use of Heckman (1979) estimation for the wage impact regressions, which uses a probit employment equation to generate an Inverse Mills ratio for each observation in a second-stage wage regression. Given certain assumptions, this removes bias from the censored wage variable (the dependent variable in second equation).

Using this approach allows us to decompose overall labor income impacts of child sponsorship into the impact from formerly sponsored children obtaining employment and the impact on wages conditional on employment. These two effects are seen by differentiating the expected average wage, $E(w)$, where

$$E(w) = \Phi(\mathbf{z}'\boldsymbol{\gamma}) \cdot w(\mathbf{x}'\boldsymbol{\beta}) \tag{2.2}$$

and $\Phi(\mathbf{z}'\boldsymbol{\gamma})$ is the probability that an individual works and earns a wage, based on characteristics $\mathbf{z}$, and $w(\mathbf{x}'\boldsymbol{\beta})$ is the individual's wage, conditional on working, based on characteristics $\mathbf{x}$. To estimate $\boldsymbol{\beta}$ without assuming arbitrary functional forms, $\mathbf{z}$ should have one or more variables that are excluded from $\mathbf{x}$; we use the individual's number of children, which strongly

[7] An identical set of instruments are used in Wydick et al. (2013b).

affects the probability of employment but should have relatively little effect on wages.[8] Both $\mathbf{x}$ and $\mathbf{z}$ include the sponsorship (treatment) variable, T. Differentiating (2.2) with respect to T, and setting variables to their means, gives

$$\frac{\partial E(w)}{\partial T} = \frac{\partial \Phi(\mathbf{z}'\boldsymbol{\gamma})}{\partial T} \cdot w(\bar{\boldsymbol{x}}'\boldsymbol{\beta}) + \frac{\partial w}{\partial T} \cdot \Phi(\bar{\boldsymbol{z}}'\boldsymbol{\gamma}) \tag{2.3}$$

The first term gives the impact of sponsorship on income from its employment effect; the second is the impact of sponsorship on wages, conditional on employment.

Both terms in (2.3) are obtained using Heckman's method; $\Phi(\mathbf{z}'\boldsymbol{\gamma})$ is estimated using a probit specification, and $w(\mathbf{x}'\boldsymbol{\beta})$ is essentially equation (2.1). To calculate the standard errors of the employment effect (the first term in (2.3)), a bootstrapping procedure is used; estimates of $\frac{\partial \Phi(\mathbf{z}'\gamma)}{\partial T}$ and of the average wage are obtained from a random draw (with replacement) from the sample. These two estimates are multiplied for each bootstrap iteration and (household-level clustered) standard errors are obtained from 500 bootstrapped replications. Similarly, the impact of sponsorship on wages (the second term in (2.3)) is the product of the estimate of $\frac{\partial w}{\partial T}$ and mean labor market participation; for each bootstrapped sample (with replacement), the entire estimation procedure, which combines estimates of the probit equation with those of the wage equation, is implemented, and 500 bootstrap replications are used to obtain

[8]It is possible that there are unobserved characteristics that reduce individuals' wages and also influence their fertility (e.g. tastes for children, or lower labor productivity, which reduces the opportunity cost of raising children). This could cause the error term in the wage equation to be correlated with the number of children, invalidating number of children as an identifying variable in the selection equation. This seems unlikely for men, but it may occur for women. To check the robustness of our results, we tried two alternative approaches. First, we used "electrified household" as the excluded variable. Such households are more likely to be near labor markets, so that individuals in them are more likely to be employed, but it is unlikely that this variable directly affects wages. Second, we used no exclusion restriction and thus relied on the assumed normality of the probit error term to identify our selection term. These two approaches yielded very similar results and thus suggest that our findings are robust to different exclusion restrictions.

standard errors.

For all 10,144 individuals in the study, interviewers attempted to obtain current labor income, which in our use of the term covers fixed wages paid by an employer, itinerate wages, estimated monthly income from a small business, or income from farming; all of this we refer to as wages. We did not collect data on non-labor income (such as returns on assets), which were infrequently realized among the low-income households in our sample. For 83% of those who were reported to be working for any wage at all, they or their family members reported on the wages of the individual. For the remaining 17% no one could provide a wage figure, but for nearly all of these individuals, family members knew their completed schooling and current occupation. Using data on education, occupation, gender and age (and country fixed effects), labor income was imputed for all individuals in the sample. Two estimates of labor income impacts from sponsorship were thus implemented; one drops the 17% of the sample without wage data, and the other imputes labor income values to all individuals in the sample, including those with observed wages.[9] A hybrid in which we impute labor income only for missing observations yields estimates very similar to the latter estimates. Assuming that any imputation errors are independent of the explanatory variables in equation (2.1), estimates using imputed values are consistent and unbiased (Wooldridge, 2010, p. 77). To carry out estimations by country and gender we use the imputed labor income, which yield slightly lower (yet more precise) impact estimates than do our directly reported wage data.

To examine the impact of child sponsorship on adult wealth, we examine two broad categories: indicators of current dwelling quality, and current ownership of common consumer durable goods and land. The dwelling quality measures include the presence of an indoor toilet, electricity, walls constructed of sturdy materials (e.g. wood or concrete, rather than mud or sticks), high quality roofs (constructed from tile, concrete, or high-quality wood, rather than thatch, leaves, or low-quality corrugated iron), and high quality floors (concrete,

[9]The few individuals whose imputed wages are less than or equal to zero are assigned non-working status.

wood, or tile, rather than dirt floors or floors made from other natural materials). For the second wealth proxy, information was collected on ownership of mobile phones, bicycles, motorcycles, automobiles, and land.

To address issues of over-testing and joint-testing of related hypotheses, two types of indices were created. The first simply weights each of the five variables within a group equally; OLS and GMM IV estimations are then carried out on these simple indices. Secondly, for each of these two groups of variables we created an Anderson (2008) summary index. This index is created by de-meaning each of the dependent variables in the respective group j ($j \in$ dwelling, consumer goods), then weighting each observation by the sum of its row entries across the inverted variance-covariance matrix of the dependent variables in the group. Specifically, each observation i in group j receives a weight (index score) of $\bar{s}_{ij} = (\mathbf{1}'\Sigma^{-1}\mathbf{1})^{-1}(\mathbf{1}'\Sigma^{-1}y_{ij})$, where $\mathbf{1}$ is an $m \times 1$ column vector of 1's, Σ^{-1} is the $m \times m$ inverted covariance matrix, and y_{ij} is the $m \times 1$ vector of outcomes for individual i. Relative to the simple index, the Anderson Index gives more weight to dependent variables within the grouping that are least correlated with other variables, and hence embody the greatest degree of unique information.

2.3 Results

Table 2.2 provides summary statistics for the data. Monthly labor income is \$16.67 higher among those who were sponsored as children ($p < 0.01$). This mainly reflects a higher employment rate (54.5% to 47.9%) among formerly sponsored individuals ($p < 0.01$). This is evident in Figure 2.2; conditional on positive labor income, (log) income is only slightly higher conditional on positive income, but formerly sponsored individuals show far more positive labor income observations.[10] Figure 2.3 illustrates the program's impact in a discontinuity diagram; non-parametric estimation shows that labor income is somewhat higher for individuals in untreated (relative to treated) households among those over age 12 when

[10] To show the density of incomes equal to zero, we specify log income as log(income+1).

the program began. However, among those 12 or younger when the program rolled out, income is clearly higher in treated households.

2.3.1 Impact on Income

The impacts of child sponsorship on adult labor income by current age are best seen visually. Figure 2.4 presents non-parametric estimations of the labor market income trajectories of sponsored (upper line) and unsponsored (lower line) individuals; the impact of sponsorship appears to increase over time (bandwidth=1, Epanechnikov kernel). While differences in income are small in the twenties and thirties, they grow substantially as individuals reach their forties, beyond which our data no longer contains observations of older formerly sponsored individuals.

Heckman estimates of income impacts are provided in Table 2.3 in columns (1) through (12). The first row of Table 2.3 presents the marginal impacts of child sponsorship on the probability of working. Column (1) gives estimates without household fixed effects and omits missing income observations, column (2) adds household fixed effects, column (3) uses household fixed effects with imputed labor income (and thus includes observations with missing income data), and columns (4), (5), and (6) provide bootstrapped IV-Heckman estimates that are analogous to the estimates in columns (1), (2), and (3). These estimates are from instrumental variable (IV) estimations in which we regress treatment on our vector of instruments and controls in the first stage and then carry out the Heckman estimation in the second stage, bootstrapping clustered standard errors at the household level for the entire process with 500 replications. The first three columns yield estimated marginal effects of 0.096, 0.079, 0.068, for the first-stage (probit) estimations. The second three columns yield estimates of 0.116, 0.191, and 0.186 of the marginal effect of the program on employment. All of the former are significant at $p < 0.01$, and the latter at $p < 0.05$ and $p < 0.10$. While the IV estimates are higher than the estimates in columns (1)-(3), Hausman tests cannot reject the null hypothesis that the standard probit estimate is consistent (the lowest p-value is 0.136).

The second row of Table 2.3 provides estimates of $\frac{\partial\Phi(\mathbf{z}'\boldsymbol{\gamma})}{\partial T} \cdot w(\bar{\boldsymbol{x}}'\boldsymbol{\beta})$ in (3), the increased income from sponsorship via greater employment. These impacts range from \$12.81 per month in column (3) to \$38.00 in column (5). Because we cannot reject the consistency of the standard probit estimates (Hausman test $t = 1.49$), we emphasize the average impact in columns (1) to (3), which is \$15.23.

The third row provides second-stage (Heckman) estimates of $\frac{\partial w}{\partial T}$, the impact of sponsorship on wages *conditional on* employment.[11] Only the \$6.06 estimate in column (3) is significant ($p < 0.05$); although two of the three IV-Heckman estimates are much larger, they are very imprecise. Thus it is only over the whole sample (including observations with missing income data) that there is evidence that sponsorship raises incomes conditional upon employment, and when the income impacts in the third row are combined with the probability of employment in the fourth row, $\frac{\partial w}{\partial T} \cdot \Phi(\bar{\boldsymbol{z}}'\boldsymbol{\gamma})$, all estimates are insignificant except for the \$2.64 estimate ($p < 0.10$) in column (3), although again two of the three IV point estimates are considerably larger. Overall, these estimates are consistent with the density functions in Figure 2.2 - the main impact of child sponsorship on income is primarily via increased employment, rather than via increased wages among those already employed.

Columns (7) through (12) in Table 2.3 show impacts on income due to the increased probability of employment from child sponsorship, replicating the estimate in column (3) for each country. Impacts are highest in India (\$37.61, $p < 0.01$), Guatemala (\$27.63, $p < 0.05$), and the Philippines (\$17.01, $p < 0.10$). And although estimates are positive in every country, they are lower and statistically insignificant in Bolivia (\$8.19), Uganda (\$7.19)

[11] The log of labor income is typically used in such regressions, but in decomposing income effects from the Heckman equation between labor force entry and the marginal income increases contingent on employment, it becomes more convenient to use levels rather than logs. Results using log income are similar, but this specification becomes very cumbersome analytically and yields little benefit.

and Kenya ($1.61).[12] Joint tests for differences across continents indicate significantly lower impacts in Africa than in Asia and Latin America, likely due to comparatively low economic opportunity in these two African countries. This is true even though educational impacts in our companion paper were found to be much stronger in the African countries.

In Table 2.4 we split up our Heckman estimations by gender, and find that the monthly labor income effects from increase probability of employment are nearly identical in the (first-stage) standard probit estimations ($12.60 for men, $12.75 for women, both $p < 0.01$). (These are equivalent to our estimates in column (3) of Table 2.3.) IV estimates for both are larger, but insignificant. There is a positive impact on wages conditional on employment for men ($6.74, $p < 0.01$), but this effect is zero for women. IV estimates of the marginal effect on men's wages are very large, $33.15 and $53.08, but imprecisely measured. All effects for the impact on the marginal wage for women are low and insignificant. Thus sponsorship yields an increase in girls' future labor income of $12.75, resulting solely from greater labor market participation. But the total impact from sponsorship on boys' future labor income is $19.34–$12.60 from higher labor market participation and $6.74 from higher wages conditional upon labor market participation.

We test for income spillovers onto unsponsored siblings and other children of eligible age within program villages using a joint test of the significance of the linear combinations $[\alpha_1 - \alpha_2] - [\gamma_1 - \gamma_2]$ and $[\beta_1 - \beta_2] - [\gamma_1 - \gamma_2]$ in (2), but find no evidence of either positive or negative spillover effects in either case ($p = 0.987$, 0.195, respectively). Thus we conclude

[12] Other estimation models yield similar estimates. For example, Ordinary Least Squares (OLS) estimates for the full sample, including those who are not working, and over a variety of specifications, yield significant ($p < 0.01$) estimates ranging from $16.60 to $19.05. While these OLS estimates capture both employment selection and marginal wage effects, they may be biased because they omit the Inverse Mills ratio, which is included in the Heckman estimation as a right-hand-side regressor. Tobit estimates, which are not preferred because they assume the impacts of the explanatory variables to be the same for the selection and marginal wage effects, yield significant ($p < 0.01$) estimates between $12.53 and $24.51.

that the benefits of international sponsorship on adult income appear to be limited to the sponsored child.

Figure 2.5 shows that the impacts on income appear to be smallest among both the least educated and most educated mothers of sponsored children. The largest difference in income between sponsored and non-sponsored children occurs among children of mothers with a primary school education, which is perhaps enough education to offer complementary support to the sponsorship program but not so much that the counterfactual levels of education for children would be high even without sponsorship.

Because most of our impact on income is on the extensive margin (labor market participation) rather than the intensive margin (higher wages), one possibility is that child sponsorship simply encouraged individuals on the margin of labor market participation to move into the labor market. Perhaps the sponsorship program merely raised aspirations for labor market activity rather than genuinely increasing the returns to labor market participation. In this case, the income gains we show here might be only slightly higher than the opportunity cost from non-wage work such as raising children or subsistence farming.

One way to test whether the income gains from child sponsorship truly increased income or simply substituted income for non-wage opportunity costs of similar value is to estimate the increase in labor market value of sponsorship via its impact on greater schooling completion. In this sense we estimate the individual terms of

$$\frac{\partial E(w)}{\partial T} = \left[\frac{\partial \Phi(\mathbf{z}'\gamma)}{\partial s} \cdot w(\bar{\boldsymbol{x}}'\boldsymbol{\beta}) + \frac{\partial w}{\partial s} \cdot \Phi(\bar{\boldsymbol{z}}'\boldsymbol{\gamma})\right] \frac{\partial s}{\partial T'} \tag{2.4}$$

where s is schooling, which carries out the same estimation except measuring impacts via the impact of the program on added schooling completion.

Schooling exhibits significant impacts on labor market participation and the number of children an individual has as an adult. Table 2.5 shows that each additional year of schooling increases labor market participation by 2.3 percentage points overall, 0.032 for women and 0.014 for men. It also reduces the number of children (as shown in negative binomial estimations) in columns (4), (5), and (6) by -0.094 children for women per added year of education and by -0.054 for men.

To check whether the program truly increased the financial returns to labor market participation via its impact on schooling we present results from a three-stage procedure in Table 2.6. In this table we jointly estimate the impact of child sponsorship on schooling (essentially replicating the results in our companion paper) through both OLS and GMM IV estimation. The second stage estimates a Heckman wage equation on the non-treated individuals in our sample as a function of total years of education and covariates, in which we obtain an estimate of the monthly income gains from an added year of schooling on the non-treated. In the third stage, we multiply the impact coefficients from the first and second stages by the expected monthly wage, *i.e.* (sponsorship program impact on schooling) $\times$ (schooling impact on labor market participation) $\times$ (mean monthly labor market wage), to obtain the mean impact of child sponsorship on wages via its impact on schooling and the added value that additional years of schooling yield in the labor market, bootstrapping the entire process with 500 replications. These estimates of the impact of the sponsorship program via the effects of education on labor market participation are given in the middle set of estimates in Table 2.6, and they show an increase in monthly wages for men between \$3.54 and \$6.56; the estimates for women lie between \$3.83 and \$6.21.

The third row of estimations repeats this exercise for the income impact of the program via the higher wages the added schooling from the program yields conditional upon employment, *i.e.* (sponsorship program impact on schooling) $\times$ (schooling impact on marginal wage, conditional on employment) $\times$ (probability of employment). Estimates here range from \$2.36 to \$3.92 for men and from \$4.11 to \$8.37 for women. The bottom row of the table shows the total income impacts that accrue to sponsored children simply via added education, \$8.04 to \$10.48 per month overall, a little higher for women, thus representing perhaps 2/3 of the impact of the program. Notably, however, these estimates indicate a higher impact for women on the marginal wage conditional on employment than is actually realized in the program. This may be because young women who finish the program show high rates of entry into the labor market, but choose to undertake relatively low-paid positions commensurate with the human service perspective of the program rather than a focus on higher adult incomes, where this message may have taken root more deeply among formerly sponsored women than among formerly sponsored men.

In many instances the decision to enter the labor force is commensurate with other demographic decisions regarding, for example, marriage and childbearing. Figure 2.6 shows kernel densities of the number of children in the families of formerly sponsored and unsponsored individuals, now adults, where the diagram seems to indicate smaller families in adulthood among the formerly sponsored. Although Figure 2.7 does not appear to show substantial differences in marriage rates over age by sponsorship status for either gender, Table 2.7 presents modest evidence that sponsorship may slightly reduce the probability of marriage. Whereas in other estimations spillover effects to siblings were found to be insignificant, and so were omitted from our regression tables, here we find evidence of significant spillovers and therefore present results that account for spillover effects onto (younger) siblings. We find very mild evidence that sponsorship causes a roughly 7 percentage point reduction in the probability of marriage at the time of the survey ($p < 0.10$) with effects reasonably uniform across age groups, perhaps having a slightly larger impact on marriage from age 17 to 21, although this is measured with noise.

In the negative binomial estimations in columns (4) to (6) of Table 2.7 we see sponsorship yielding a small reduction in family size. This table presents negative binomial estimation coefficients, but these magnitudes translate to marginal effects (accounting for spillovers) of 0.248 children overall, and 0.64 fewer children for older sponsored children age 40-45 at the time of the survey, with lower point estimates for the younger adults. The impacts on childbearing are likely greater on older individuals simply because birth rates were much higher when these individuals were sponsored back in the early 1980s, and so the impact of greater labor force entry has a stronger effect on birth rates. This can be seen in Figure 2.8, for example, where birth rates are only slightly lower for the formerly sponsored, but then fall considerably among the older cohort. We do not present impacts by gender; they are virtually identical between men and women.

The increase in income from child sponsorship occurs through different career choices of formerly sponsored men and women. Table 2.8 presents multinomial logit estimates of the impacts of sponsorship on vocational trajectories, separately by gender. Sponsorship of boys leads them in adulthood into two main types of jobs: K-12 teachers and employees in lower-skill technology jobs, such as work in call centers, with some evidence of an increase in blue-collar employment. Formerly sponsored men are roughly 62% more likely to be teachers in adulthood relative to the counterfactual and are about 44% more likely to have a semi-skilled technology job or work in a call center. Sponsorship of girls makes them 50% less likely to be involved in agriculture as adults, 55% more likely to have a clerical job, 60% more likely to work in finance or for a large corporation, 148% more likely to have a semi-skilled technology job (starting from a very low baseline), and 93% more likely to be a nurse or health professional.

2.3.2 Wealth Impacts

Finally, we consider the impact of child sponsorship on indicators of wealth in adulthood. Results show that individuals who were sponsored as children live in better houses as adults. Both the simple index and the Anderson index indicate significant impacts of sponsorship on adult dwelling quality. Specifically, OLS (linear probability model) household fixed-effect estimates in Table 2.9 indicate that sponsorship increases the probability that a home has electricity by 2.9 percentage points, raises the probability of having improved walls by 2.5

percentage points, and increases having improved floors by 1.9 percentage points. GMM-IV estimates are smaller and insignificant for specific improvements, but larger and significant for both dwelling indices.

Does child sponsorship increase consumer durable ownership in adulthood? The only asset with a statistically significant effect is the probability of owning a mobile phone, an increase of 5.4 percentage points in the OLS estimate and 18.3 percentage points in the IV estimate (baseline of 76.8%). We find no evidence that sponsorship increased ownership of bicycles, motorcycles, vehicles, or land; the IV coefficients on both consumer good summary indices are insignificant.

Tests for household and village level spillovers show no significant effects on aggregated dwelling indicators ($p = 0.237$ and $p = 0.523$, respectively) or consumer durables ($p = 0.333$ and $p = 0.536$). Although our research on educational impacts provides evidence for spillovers onto younger siblings, particularly in secondary school completion (Wydick, Glewwe and Rutledge, 2013b), we find no evidence of income or wealth spillovers in our data.

Tables 2.10 and 2.11 disaggregate wealth impact estimations by gender. Not surprisingly, impacts on dwelling quality appear to be higher for formerly sponsored men than for women. OLS estimates for men are positive on every dwelling category–indoor toilet, electrification of the household, improved walls, improved roof, and improved floor–and statistically significant for every category except improved walls. Both the simple dwelling index and the Anderson dwelling index are also strongly significant. However, IV estimates for men are all insignificant. OLS estimates for women indicate that sponsorship appears to affect only the probability of living in a home with electricity, and even this is not significant for the IV estimates. These differences in effects by gender presumably reflect the larger overall impact of sponsorship on the incomes of formerly sponsored men, and they may also reflect that a husband's income has a stronger influence on the type of dwelling in adulthood.

Impacts on consumer durables are virtually identical between formerly sponsored men and women, where estimations indicate a strong and significant impact on cell phone ownership for both genders, but no effect on ownership of a bicycle, motorcycle, car, or land.

Another measure of wealth in adulthood relates to where an individual resides as an adult. Does a grown adult remain in her parents' home, in a rented home, or in an owned home? Typically living in a home apart from parents is desirable after a certain age, especially for married couples, but this is not always economically feasible. Baseline values among the untreated show 46.8% living in the parents' home, 23.8% living in a rented home, and 29.4% living in a home owned by the individual or jointly owned with a spouse. Table 2.12 shows multinomial logit estimations indicating that formerly sponsored individuals are much less likely to remain living in their parents' home. There is some indication that women are more likely to live in an owned home, but marginal effects are insignificant. Instead of living in their parents' home as adults at the time of the survey, sponsored individuals are 4.8 percentage points more likely to live in a home rented themselves, about two percentage point higher for men than for women.

2.4 Conclusion

International child sponsorship is a leading form of individual contact and financial assistance between ordinary people in developed countries and the poor in developing countries, yet little has been known about the impact of these programs on the economic outcomes in adulthood of sponsored children. Our more conservative Heckman estimates from a six-country study of 10,144 individuals show that child sponsorship is responsible for increases in monthly income of about \$15 over an unconditional baseline of \$75, or an increase of approximately 20%. This effect of child sponsorship on future labor income is due principally to sponsored children entering the labor market as adults who would not have done so otherwise, particularly for women. We find that men realize an added \$6.74 of monthly income from higher wages conditional upon employment, but that estimations on the impact on wages of women cannot reject a null hypothesis of zero.[13]

Given that the cost of sponsorship to sponsors was \$28 per month during the time in the 1980s and 1990s when the individuals in our study were sponsored, and that the average length of sponsorship was 9.3 years, a monthly income increase of a little over \$15 over an average lifetime of work implies a modest financial rate of return to child sponsorship of 4.5%. Note that this calculation omits both the opportunity cost of children's time in the program and other non-pecuniary returns to education in adulthood. Because the latter is likely to exceed the former, this return almost certainly represents a lower bound on the rate of return to child sponsorship.

Our estimations of impacts on wealth in adulthood find significant impacts on adult dwelling quality from child sponsorship on proxies for adult wealth, where we find that sponsored children - especially males - are more likely as adults to live in better housing, homes with electricity, and with roof and floor made of superior construction. Formerly sponsored men are more likely to live in homes with indoor plumbing. Impacts on adult consumer good ownership, however, are more modest and appear to be limited to substantially greater ownership of mobile phones among both formerly sponsored men and women. We also present (modest) evidence suggesting that sponsored children have fewer children,

---

[13] As pointed out by a referee, income increases due to increased labor market participation do not account for the implicit cost of reduced leisure time, and since at the margin those two uses of time should be approximately equal the net benefit of child sponsorship via increased labor market participation may be close to zero. However, this ignores the fact that sponsorship can have a variety of non-economic effects that make leisure time more valuable, such as improved health and overall greater psychological well-being. While we cannot estimate those effects with our data, they are likely to be positive, and so even if all of the increased income is due solely to increased labor force participation the net benefit of the program is unlikely to be zero. In addition, some of our estimates in Tables 2.3, 2.4, and 2.6 find positive impacts on wages from sponsorship, and these can serve as a lower bound of the "net" impact of sponsorship on income.

along with stronger evidence that both formerly sponsored men and women are less likely to live with their parents in adulthood.

What about child sponsorship, in particular Compassion's approach to child sponsorship, could be responsible for these significant effects on income and wealth in adulthood? In related research using a separate sample of *currently* sponsored children we explore the hypothesis that child sponsorship may improve adult incomes not merely through relieving *external* constraints that improve schooling access, nutrition, and health, but through addressing *internal* constraints related to imparting a greater level of hopefulness about the future and instilling greater aspirations for schooling and adult vocation. Using data on currently sponsored children, we find in Glewwe, Ross and Wydick (2014) a causal link between child sponsorship and elevated educational and vocational aspirations among children in Kenya, and higher levels of happiness, self-efficacy, and hopefulness based on a quantitative analysis of children's self-portrait drawings in Indonesia. Although it is yet impossible to definitively identify these increased aspirations as a causal channel to the positive impact from sponsorship on income and wealth we find in this study, what is clear from our three pieces of research on child sponsorship is that child sponsorship increases aspirations and that child sponsorship also improves adult economic outcomes. We present a diagram in Figure 2.9 of what appears to us to be the causal channel for the effects we observe from child sponsorship on income and wealth in adulthood.

Most conditional and unconditional cash transfer programs, and many - if not most - educational interventions, do not seek to directly address internal constraints of children, which are also related to the fostering of non-cognitive (socio-emotional) skills. Our findings on the impacts of child sponsorship suggest that this may constitute a missed opportunity. Taken together, our results suggest that development interventions that relieve tangible external constraints, while simultaneously addressing the internal constraints faced by the poor, may realize stronger impacts than those that address external constraints alone, thus providing a basis for experimenting with new programs that embody these joint characteristics and for important future research.

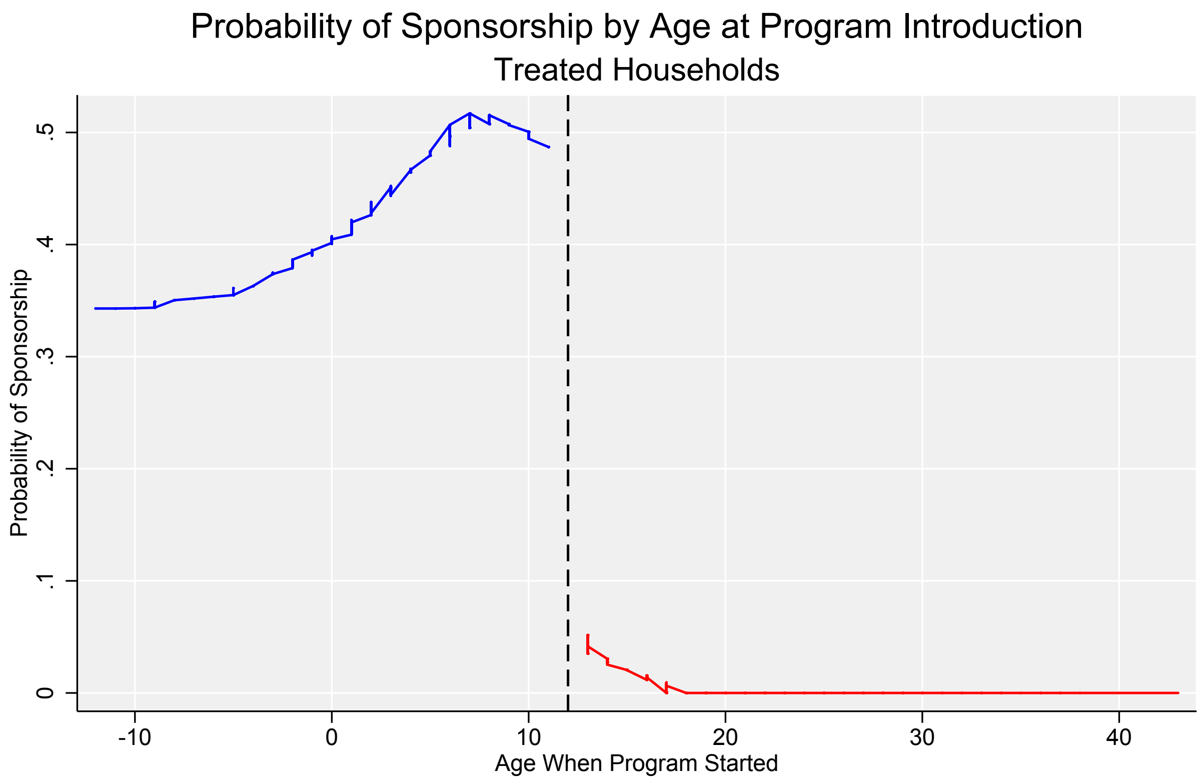

Figure 2.1: Sponsorship as a Function of Age When Program Started

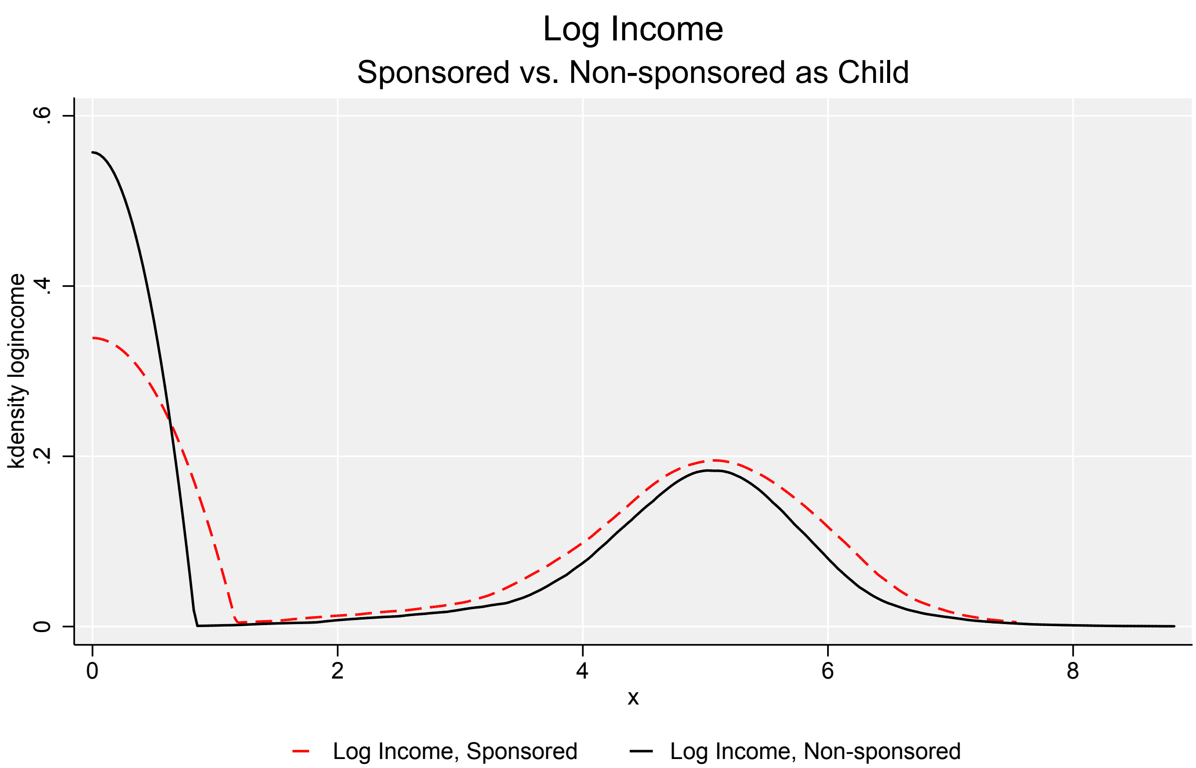

Figure 2.2: Differences in Log Income, Sponsored vs. Unsponsored

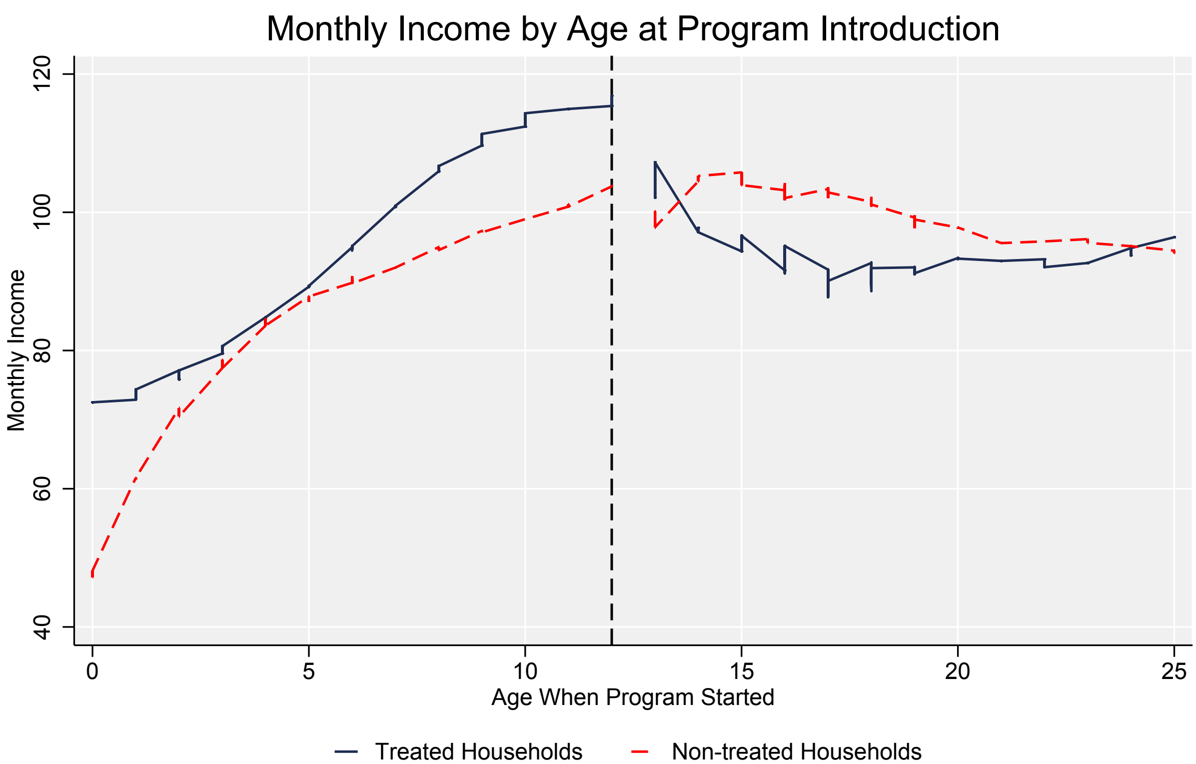

Figure 2.3: Monthly Income as a Function of Eligibility

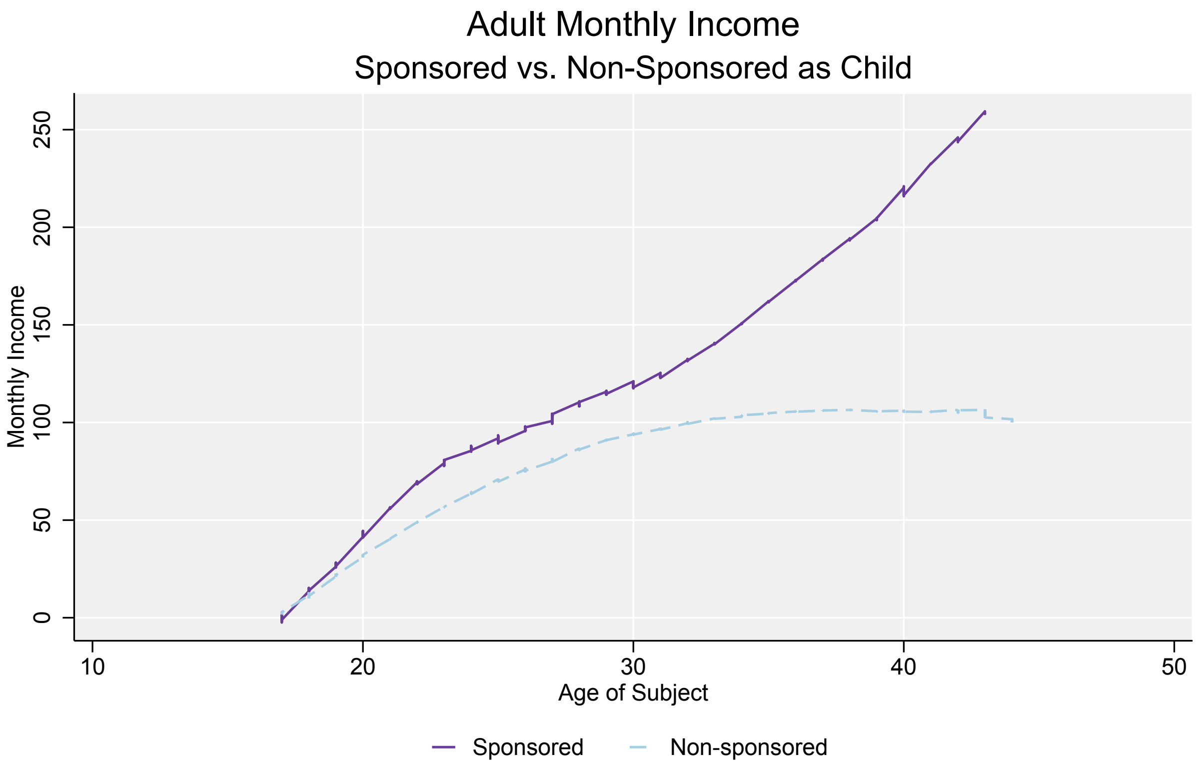

Figure 2.4: Growth in Labor Income Gap Over Time, Sponsored vs. Non-Sponsored

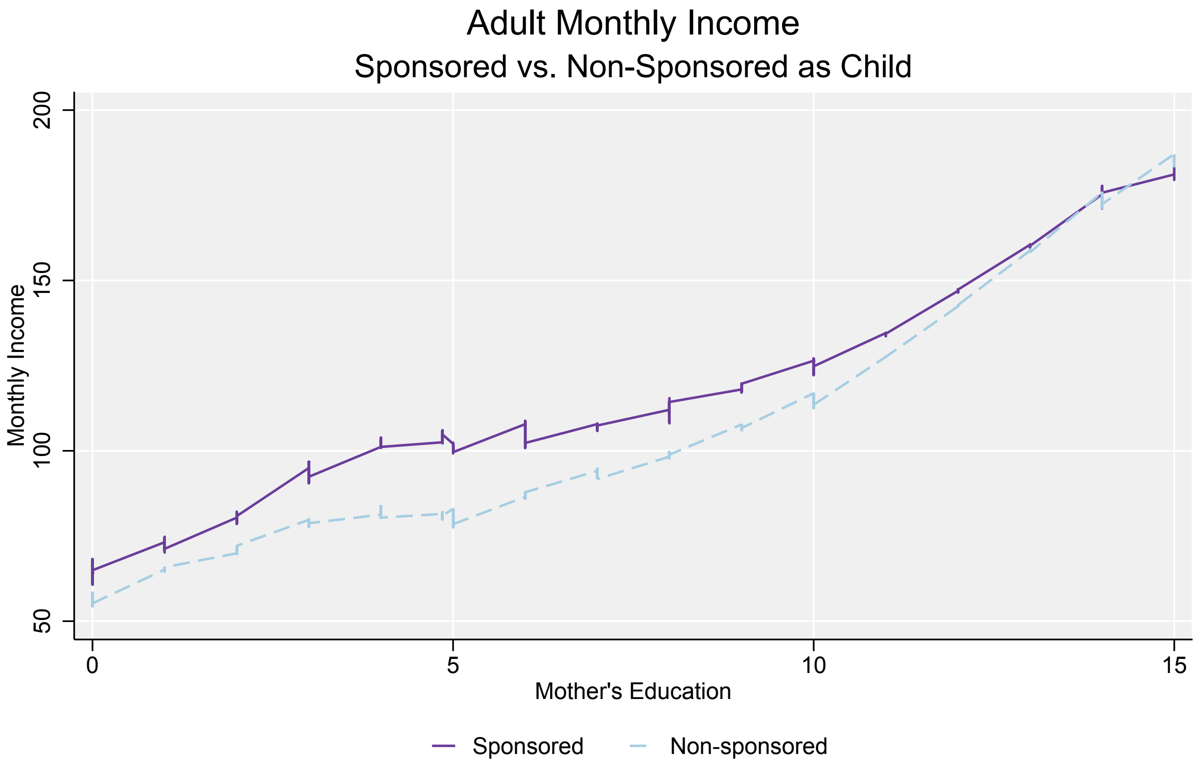

Figure 2.5: Impact on Labor Income between Sponsored and Non-sponsored by Mother's Education (Bandwidth=0.5)

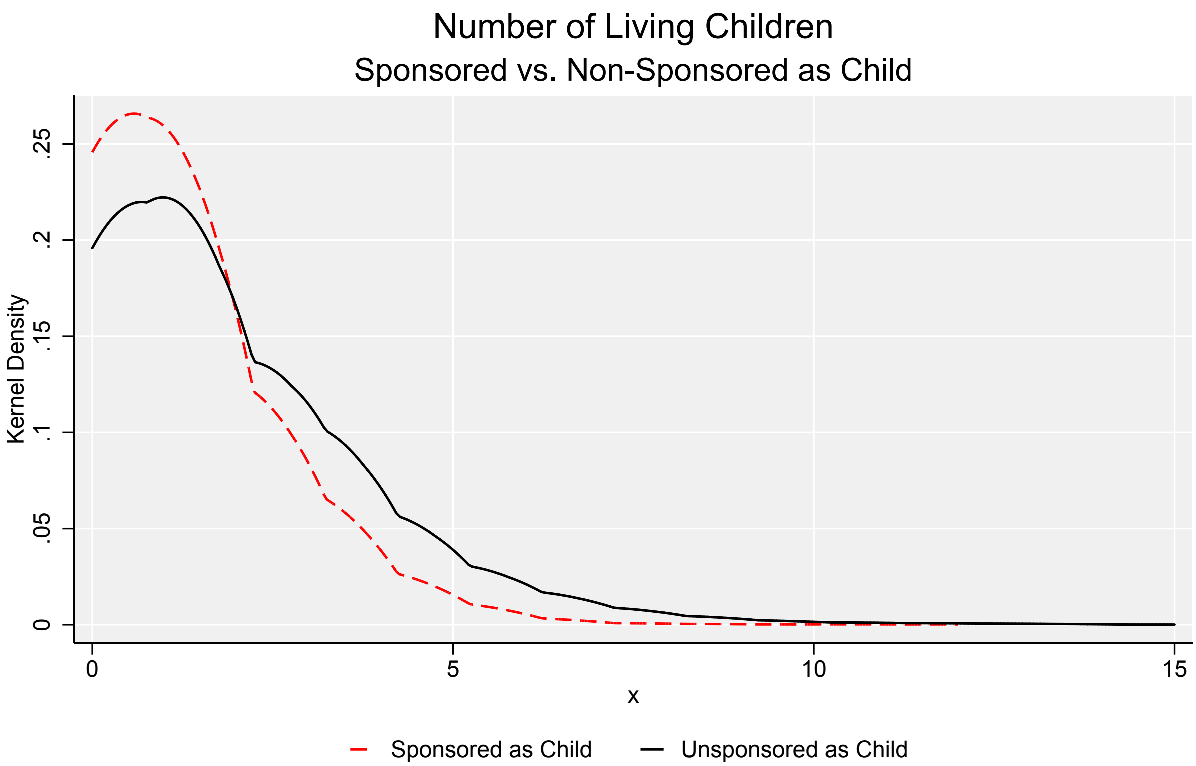

Figure 2.6: Kernel Density of Number of Children by Sponsorship Status

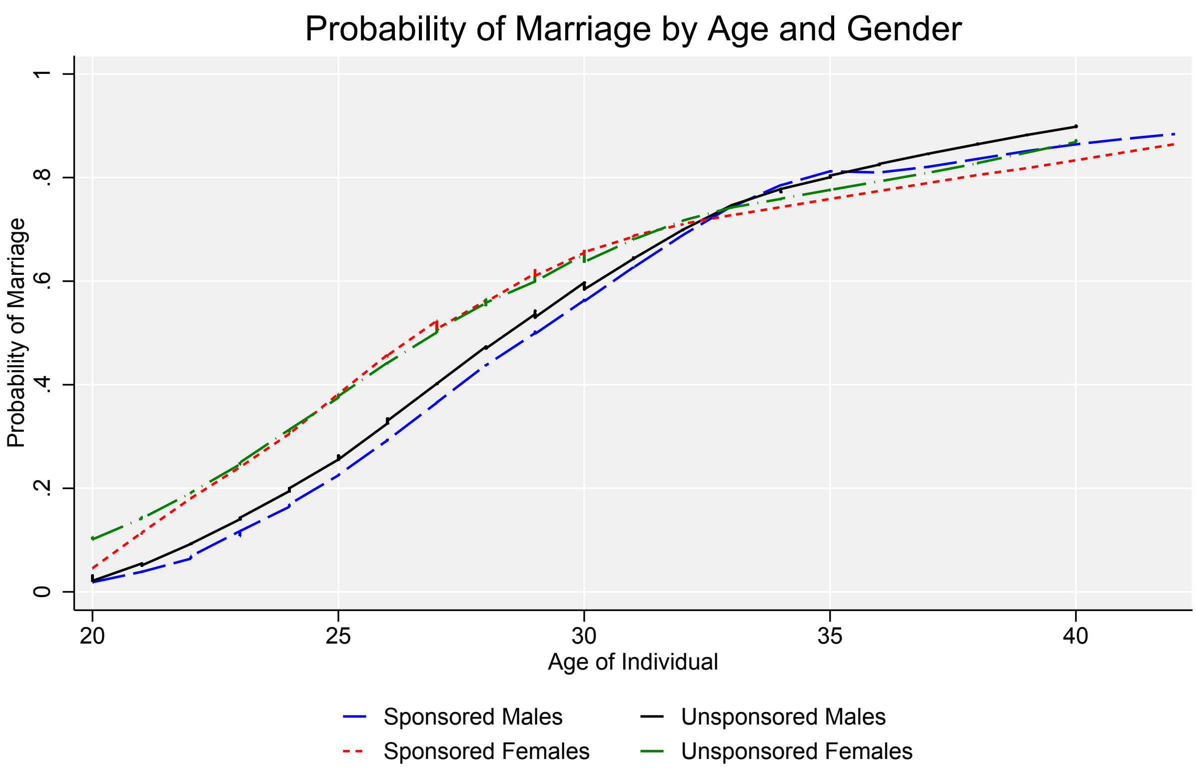

Figure 2.7: Probability of Marriage by Sponsorship, Sex, and Age

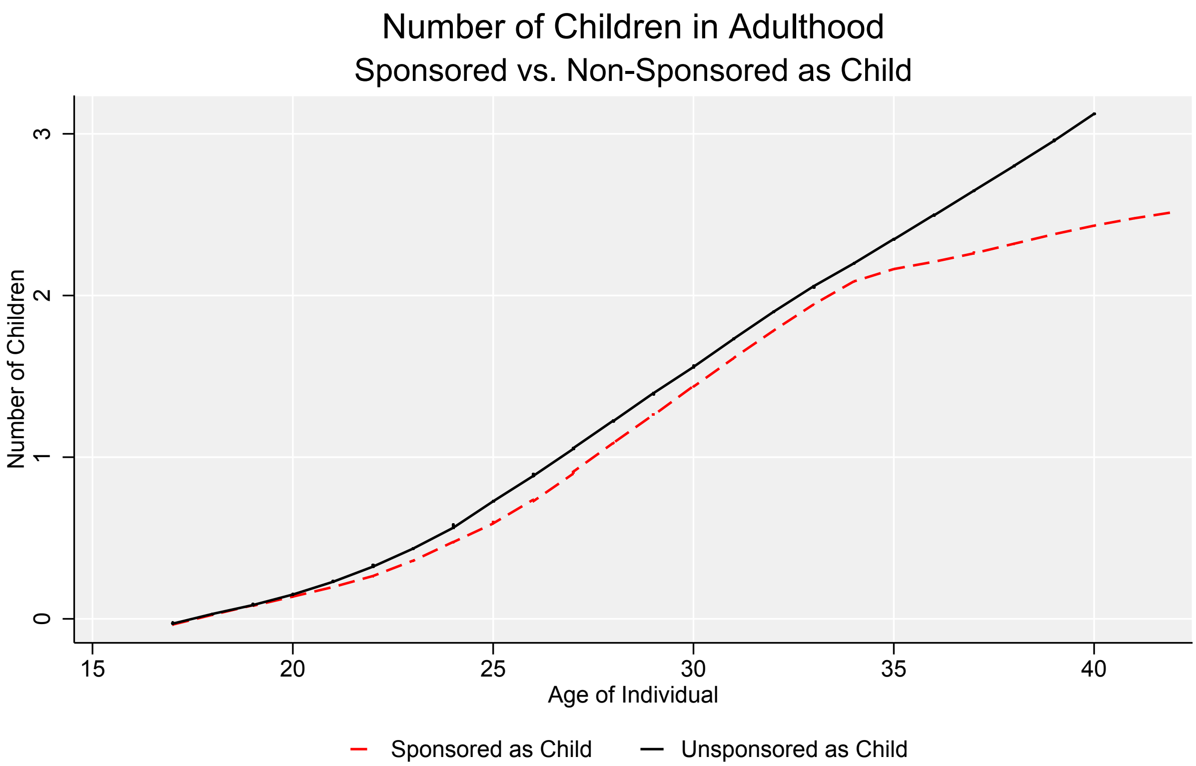

Figure 2.8: Number of Children in Adulthood by Age

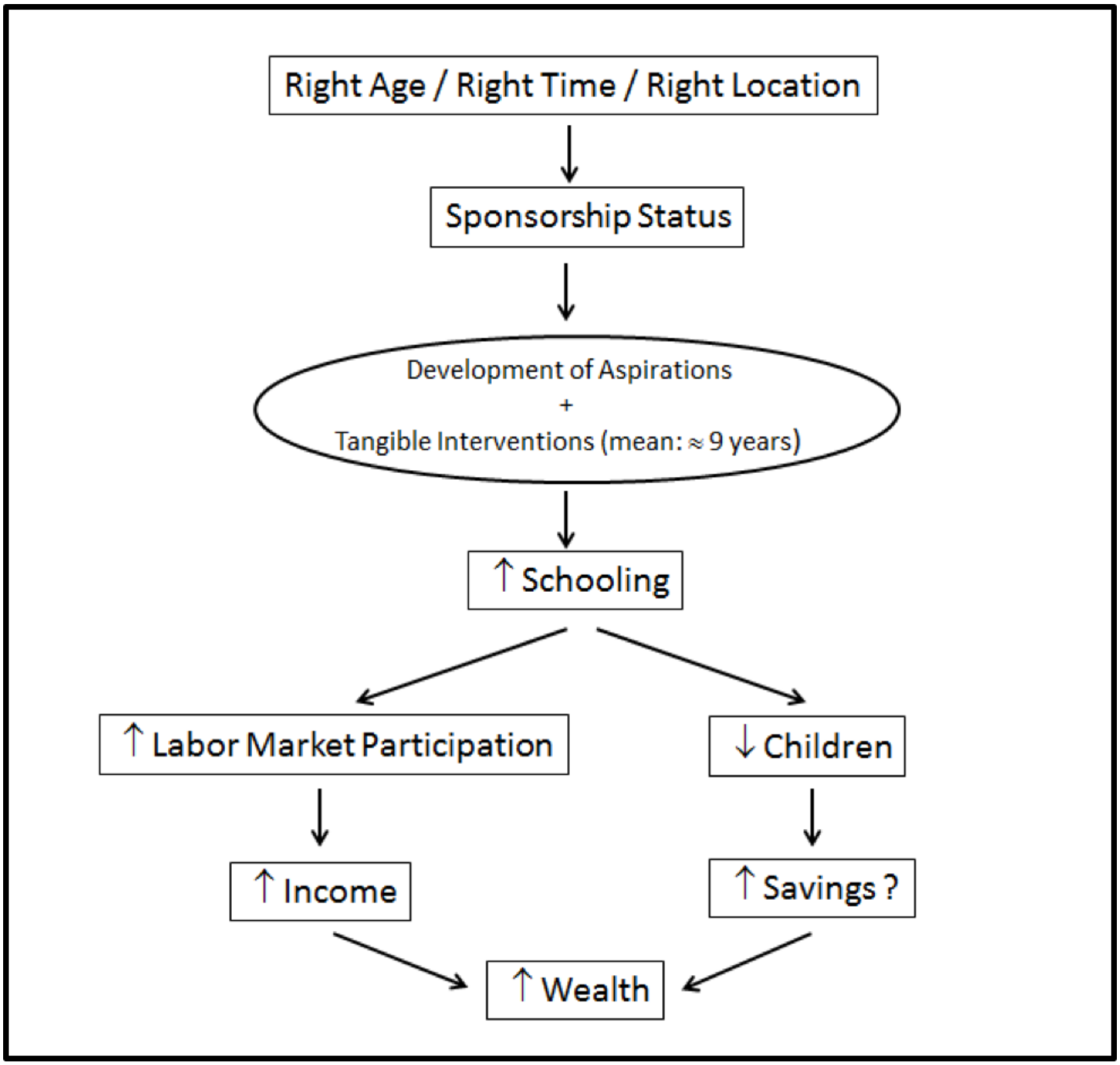

Figure 2.9: Plausible Chain of Causal Effects from Sponsorship

Table 2.1: Years of Sponsorship in Program by Country

	Mean	Std. Dev.	Frequency
Uganda	11.02	3.53	188
Guatemala	6.64	2.55	357
Philippines	7.12	4.79	237
India	10.86	3.53	221
Kenya	10.13	3.43	543
Bolivia	9.48	3.78	288
Total	**9.30**	**3.93**	**1,834**

Note: weighted mean is presented

Table 2.2: Summary Statistics

Variable	Full Sample	Sponsored	Unsponsored	Difference	p-value
Age	29.82	26.51	30.56	4.05	0.000
Sex	0.504	0.481	0.509	-0.028	0.013
Birth Order	3.041	3.006	3.049	-0.043	0.389
Mothers Educ. (Years)	4.85	4.94	4.83	0.11	0.298
Uganda	0.08	0.102	0.075	0.027	0.001
Guatemala	0.168	0.192	0.162	0.030	0.002
Philippines	0.141	0.129	0.143	-0.014	0.122
India	0.159	0.119	0.168	-0.049	0.000
Kenya	0.304	0.296	0.306	-0.010	0.418
Bolivia	0.145	0.158	0.142	0.016	0.089
Working=1	0.491	0.545	0.479	0.066	0.000
Monthly Income ($US)	77.96	91.53	74.86	16.67	0.008
Monthly Income ($US), Working=1	198.13	194.25	199.24	-4.99	0.649
Monthly Income (Imputed, $US)	90.63	104.13	87.62	16.51	0.000
Monthly Income (Imp, $US), Working=1	170.25	177.48	168.43	9.05	0.009
Housing Quality Index (Simple)	2.81	2.88	2.80	0.08	0.014
Consumer Good Index (Simple)	1.26	1.25	1.27	-0.02	0.378
Sample Size	**10,011**	**1,819**	**8,192**		

Table 2.3: Impact on Monthly Labor Income: Heckman Estimates

	(1)	(2)	(3)	(4)	(5)	(6)
	Heckman Misg. Omt. No FE	Heckman Misg. Omt. HH FE	Heckman Obs. Impt. HH FE	IV-Heckman Misg. Omt. No FE	IV-Heckman Misg. Omt. HH FE	IV-Heckman Obs. Impt. HH FE
Heckman Selection $= \frac{\partial \Phi(z'\gamma)}{\partial T}$	0.096***	0.079***	0.068***	0.116*	0.191**	0.186**
	(0.017)	(0.016)	(0.014)	(0.066)	(0.090)	(0.080)
Selection Impact on Income $= \frac{\partial \Phi(z'\gamma)}{\partial T} \cdot w(\bar{x}'\beta)$	**$17.25*** **	**$15.63*** **	**$12.81*** **	**$23.65* **	**$38.00** **	**$35.31** **
	(3.12)	(3.08)	(2.71)	(13.22)	(17.76)	(15.33)
Marginal wage Impact $\mid w > 0$ $= \frac{\partial w}{\partial T}$	-$1.17	-$5.39	$6.06**	$40.16	$48.99	$10.23
	(9.85)	(10.79)	(3.10)	(50.05)	(50.79)	(15.21)
Marginal wage Impact on Income $= \frac{\partial w}{\partial T} \cdot \Phi(\bar{z}'\gamma)$	-$0.46	-$2.12	$2.64*	$15.80	$19.28	$4.46
	(4.02)	(4.25)	(1.38)	(23.84)	(19.76)	(6.69)
Lambda	-18.85***	-56.12	-10.55***	-19.37**	-57.80	-10.31
	(3.94)	(41.19)	(3.60)	(8.03)	(52.76)	(7.50)
Observations	8,389	8,389	10,004	8,389	8,389	10,004
Mean w, untreated	$74.86	$74.86	$74.86	$74.86	$74.86	$74.86
	(196.07)	(196.07)	(196.07)	(196.07)	(196.07)	(196.07)
Mean $w \mid w > 0$, untreated	$199.24	$199.24	$199.24	$199.24	$199.24	$199.24
	(278.47)	(278.47)	(278.47)	(278.47)	(278.47)	(278.47)
BY COUNTRY:	(7)	(8)	(9)	(10)	(11)	(12)
Selection Impact on Income $= \frac{\partial \Phi(z'\gamma)}{\partial T} \cdot w(\bar{x}'\beta)$	Col. (3) OLS HH FE **Uganda**	Col. (3) OLS HH FE **Guatemala**	Col. (3) OLS HH FE **Philippines**	Col. (3) OLS HH FE **India**	Col. (3) OLS HH FE **Kenya**	Col. (3) OLS HH FE **Bolivia**
Sponsored	**$7.19**	**$27.63** **	**$17.01* **	**$37.61*** **	**$1.61**	**$8.19**
	(7.82)	(8.35)	(9.54)	(6.47)	(3.57)	(6.09)
Observations	809	1,680	1,407	1,599	3,051	1,458
Mean w, untreated	$36.90	$56.65	$115.33	$131.96	$31.22	$57.73
	(144.30)	(104.66)	(253.02)	(167.22)	(89.62)	(120.46)
Mean $w \mid w > 0$, untreated	$154.99	$193.34	$301.78	$198.28	$111.62	$165.36
	(264.39)	(104.68)	(333.85)	(169.90)	(140.65)	(154.36)

Heckman estimations include controls for age, gender, sibling order, and oldest sibling. Selection on Impact multiplies marginal effect of first-stage tobit by $E[w \mid w > 0]$ for sample ($178.07). First-stage F-test for instrumental variable estimation yields $F = 225.5$ ($p < 0.001$). Hausman test for efficiency of Heckman vs. IV Heckman fails to reject null of non-instrumented Heckman efficiency ($p = 0.1363$).

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$

Table 2.4: Impact on Monthly Labor Income by Gender: Heckman Estimates

	(1) Heckman All Obs Imp. HH FE Men	(2) IV-Heckman Misg. Omt. No FE Men	(3) IV-Heckman Misg. Omt. HH FE Men	(4) Heckman All Obs Imp. HH FE Women	(5) IV-Heckman Misg. Omt. No FE Women	(6) IV-Heckman Misg. Omt. HH FE Women
Heckman Selection $= \frac{\partial \Phi(z'\gamma)}{\partial T}$	0.063***	0.104	0.258	0.073***	0.093	0.138
	(0.023)	(0.134)	(0.205)	(0.022)	(0.071)	(0.106)
Selection Impact on Income $= \frac{\partial \Phi(z'\gamma)}{\partial T} \cdot w(\bar{x}'\beta)$	**\$12.60*** **	**\$22.03**	**\$51.78**	**\$12.75*** **	**\$18.07**	**\$26.50**
	(4.44)	(27.60)	(41.95)	(3.90)	(13.50)	(20.01)
Marginal wage Impact | $w > 0$ $= \frac{\partial w}{\partial T}$	\$12.90***	\$69.44	\$111.18	\$0.43	-\$3.77	\$20.97
	(4.14)	(83.70)	(87.69)	(4.62)	(72.58)	(54.89)
Marginal wage Impact on Income $= \frac{\partial w}{\partial T} \cdot \Phi(z'\gamma)$	**\$6.74*** **	**\$33.16**	**\$53.08**	**\$0.15**	**-\$1.16**	**\$6.48**
	(2.21)	(41.36)	(41.57)	(1.65)	(21.35)	(16.72)
Lambda	-5.41***	-13.48***	-109.50	-35.90***	-27.05	-42.37
	(1.71)	(4.95)	(24.79)	(6.19)	(50.32)	(101.10)
Observations	5,048	4,197	4,197	4,956	4,192	4,956
Mean w, untreated	\$100.70	\$96.98	\$96.98	\$47.54	\$58.83	\$58.83
	(185.69)	(195.92)	(195.92)	(114.82)	(183.47)	(183.47)
Mean $w \mid w > 0$, untreated	\$201.89	\$203.12	\$203.12	\$175.02	\$190.29	\$190.29
	(220.71)	(242.57)	(242.57)	(162.01)	(289.65)	(289.65)

Heckman estimations include controls for age, gender, sibling order, and oldest sibling. Selection on Impact multiplies marginal effect of first-stage tobit by $E[w \mid w > 0]$ for sample by gender. first-stage F-tests for instrumental variable estimation yields $F = 110, 116$, respectively for Men and Women ($p < 0.001$).
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$

Table 2.5: Schooling Impacts on Labor Market Participation and Number of Children

	Labor Market Participation			Number of Children		
	(1)	(2)	(3)	(4)	(5)	(6)
Variables	All	Men	Women	All	Men	Women
	OLS, Household Fixed Effects			Negative Binomial Regression		
Schooling	0.023***	0.014***	0.032***	-0.074***	-0.054***	-0.094***
	(0.002)	(0.004)	(0.003)	(0.005)	(0.007)	(0.007)
Observations	8,348	4,180	4,168	9,955	5,025	4,930
R-squared	0.045	0.034	0.032	0.303	0.298	0.303

Regressions include controls for age, gender, and sibling order. OLS estimations incorporate household fixed effects, but negative binomial regressions do not incorporate household fixed effects due to integer constraints. Alphas are significant at $p < 0.01$ in all negative binomial estimations. Clustered standard errors at the household level are in parentheses.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$

Table 2.6: Estimated Program Impact on Income via Labor Market Effects from Added Schooling

Variables	(1) OLS Fixed Effects	(2) IV Fixed Effects	(3) Men: OLS Fixed Effects	(4) Men: IV Fixed Effects	(5) Women: OLS Fixed Effects	(6) Women: IV Fixed Effects
Dep. Variable: Years of Schooling[a]						
Sponsored	1.11***	1.45***	1.14***	2.00***	1.09***	0.695
	(1.21)	(0.406)	(1.38)	(5.16)	(0.145)	(0.487)
R-squared	0.063	0.047	0.056	0.042	0.076	0.072
Income Impact of Program via Schooling Effect on Labor Market Participation[b]	OLS × Probit × Avg. Wage Fixed Effects	IV × Probit × Avg. Wage Fixed Effects	Men: OLS × Probit × Avg. Wage, FE	Men: IV × Probit × Avg. Wage, FE	Women: OLS × Probit × Avg. Wage, FE	Women: IV × Probit × Avg. Wage, FE
	$5.03* **	**$6.56** **	**$3.54*** **	**$6.21***	**$6.01*** **	**$3.83**
	(0.743)	(2.83)	(0.902)	(3.36)	(1.09)	(3.38)
Income impact of Program via Schooling Effect on Labor Market Wage[c]	OLS × Heckman × Prob(Emp.) Fixed Effects	IV × Heckman × Prob(Emp.) Fixed Effects	Men: OLS × Heckman × Prob(Emp.) Fixed Effects	Men: IV × Heckman × Prob(Emp.) Fixed Effects	Women: OLS × Heckman × Prob(Emp.) Fixed Effects	Women: IV × Heckman × Prob(Emp.) Fixed Effects
Years of Schooling	**$3.01***	**$3.92** **	**$2.36***	**$4.11**	**$8.37***	**$5.33**
	(1.75)	(2.13)	(1.44)	(2.80)	(4.62)	(5.08)
Heckman's Lambda	-5.62**	-5.62**	11.66	11.66	-0.90***	-0.90***
	(2.44)	(2.44)	(7.55)	(7.55)	(0.24)	(0.24)
Observations	8,348	8,348	4,180	4,180	4,168	4,168
Sponsorship Impact via Schooling ↑ Earnings from Employment + ↑ Earnings from Marginal Wage	**$8.04**	**$10.48**	**$5.90**	**$10.32**	**$14.38**	**$9.16**

[a] OLS and Instrumental Variables estimates include controls for age, gender, sibling order, oldest sibling and household-level fixed effects. First-stage F-tests for instrumental variable estimation yields $F = 225.5$ (all), 111.0 (Men), 116.7 (Women) respectively ($p < 0.001$).
[b] Reported coefficients stem from product of (sponsorship program impact on schooling) × (schooling impact on labor market participation) × (Mean Labor Market Wage). Joint estimates and include controls for age, gender, sibling order, oldest sibling and house-hold level fixed effects. Coefficients are estimated jointy; bootstrapped standard errors clustered at the household level (500 replications).
[c] Reported coefficients stem from product of (sponsorship program impact on schooling) × (schooling impact on marginal wage, conditional on employment) × (Probability of Employment). Joint estimates and include controls for age, gender, sibling order, oldest sibling and household-level fixed effects. Coeffficients are estimated jointly; bootstrapped standard errors clustered at the household level (500 replications).
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$

Table 2.7: Sponsorship Impacts on Marriage and Number of Children

	Prob. Married			Number of Children		
	(1)	(2)	(3)	(4)	(5)	(6)
	LP Model	LP w/ FE,	GMM-IV FE	Neg. Binomial	Neg. Binomial	Neg. Bin. IV
Variables	HH FE	Interact.	w/ Interact.		w/ Interact.	w/ Interact.
OLS, Household FE:						
Sponsored	-0.029*	-0.038	-1.431	-0.018	0.157	1.074***
	(0.016)	(0.094)	(1.085)	(0.038)	(0.230)	(0.422)
Sponsored + (Spon × Age 17-21)		-0.042	-0.032			
		(0.045)	(0.045)			
Sponsored + (Spon × Age 22-27)		0.001	-0.012		-0.077	0.580***
		(0.022)	(0.023)		(0.065)	(0.199)
Sponsored + (Spon × Age 28-33)		-0.012	-0.030		-0.010	0.562***
		(0.024)	(0.027)		(0.047)	(0.175)
Sponsored + (Spon × Age 34-39)		-0.021	-0.077		0.005	0.292
		(0.041)	(0.058)		(0.085)	(0.191)
Sponsored + (Spon × Age 40-45)					-0.256*	-0.265
					(0.145)	(0.307)
Intra-Household Spillovers	-0.041	-0.029	-0.001	-0.152**	-0.152***	-0.388***
$(\alpha_1 - \alpha_2) - (\gamma_1 - \gamma_2)$	(0.039)	(0.039)	(0.043)	(0.069)	(0.065)	(0.105)
Program Impact						
with Household Spillovers						
$\tau + (\alpha_1 - \alpha_2) - (\gamma_1 - \gamma_2)$	-0.071*	-0.067	-1.430	-0.170***	0.005	0.685*
	(0.039)	(0.101)	(1.608)	(0.068)	(0.240)	(0.410)
$\tau + (\alpha_1 - \alpha_2) - (\gamma_1 - \gamma_2)+$ (Spon × Age 17-21)		-0.071	-0.031			
		(0.059)	(0.064)			
$\tau + (\alpha_1 - \alpha_2) - (\gamma_1 - \gamma_2)+$ (Spon × Age 22-27)		-0.027	-0.011		-0.229***	0.191
		(0.044)	(0.047)		(0.087)	(0.157)
$\tau + (\alpha_1 - \alpha_2) - (\gamma_1 - \gamma_2)+$ (Spon × Age 28-33)		-0.041	-0.029		-0.162**	0.174
		(0.042)	(0.042)		(0.074)	(0.119)
$\tau + (\alpha_1 - \alpha_2) - (\gamma_1 - \gamma_2)+$ (Spon × Age 34-39)		-0.050	-0.077		-0.146	-0.095
		(0.056)	(0.061)		(0.101)	(0.143)
$\tau + (\alpha_1 - \alpha_2) - (\gamma_1 - \gamma_2)+$ (Spon × Age 40-45)					-0.407***	-0.653***
					(0.156)	(0.264)
Observations	10,001	10,001	10,001	10,004	10,004	10,004
R-squared	0.108	0.144	0.125			

Regressions include controls for age, gender, and sibling order. OLS and IV estimations incorporate household fixed effects. Negative binomial regressions display coefficients (not marginal effects) and omit household fixed effects with alpha significant at 1% in all regressions (rejecting null of Poisson distribution). Impacts on "Sponsored + (Spon × Age Group)" are the sum of coefficients on Sponsored added to the coefficient of Sponsored × Age Group and joint tests of these two coefficients. GMM-IV estimations instrument for schooling using sponsorship program. First-stage instrumental variable estimations yield $F = 297.12$. Clustered standard errors at the household level in parentheses.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$

Table 2.8: Impacts on Adult Vocation by Gender: Marginal Effects, Multinomial Logit Estimations

Men

Occupational Category	MN Logit Coefficients	Marginal Effects	Baseline Untreated	Occupational Category	MN Logit Coefficients	Marginal Effects	Baseline Untreated
1 Agriculture	0.042	0.0068	**0.048**	8 Small business	0.053	-0.0052	**0.045**
	(0.23)	(0.011)			(0.25)	(0.0105)	
2 Construction,	0.497**	0.012	**0.046**	9 Ministry,	0.477	0.0026	**0.0043**
Day Labor	(0.256)	(0.009)		Pastoral	(0.484)	(0.0041)	
3 Clerical, Sales	0.290	0.005	**0.047**	10 Finance and	0.329	0.0045	**0.027**
	(0.249)	(0.009)		Large Business	(0.287)	(0.0078)	
4 Blue Collar	0.349**	0.019	**0.104**	11 Police, Army,	0.471	0.0056	**0.022**
	(0.166)	(0.015)		Security, Fire	(0.350)	(0.0064)	
5 Personal	0.390*	0.012	**0.063**	12 Professional,	0.059	-0.0025	**0.024**
Services	(0.216)	(0.011)		Doctor, Lawyer	(0.331)	(0.0072)	
6 Teaching	0.920***	0.0274***	**0.043**	13 Semi-skill Tech,	0.618**	0.012*	**0.027**
	(0.233)	(0.0079)		Call Centers	(0.288)	(0.0071)	
7 Government	0.831	0.0044	**0.0078**	14 Nursing, Health,	-0.500	-0.0026	**0.0043**
	(0.506)	(0.0033)		Hospital	(0.83)	(0.0032)	

Women

Occupational Category	MN Logit Coefficients	Marginal Effects	Baseline Untreated	Occupational Category	MN Logit Coefficients	Marginal Effects	Baseline Untreated
1 Agriculture	-0.393	-0.021**	**0.042**	8 Small Business	-0.234	-0.012	**0.028**
	(0.26)	(0.010)			(0.277)	(0.0091)	
2 Construction,	1.180*	0.0036*	**0.0049**	9 Ministry,	0.361	0.0004	**0.0012**
Day Labor	(0.657)	(0.0021)		Pastoral	(1.12)	(0.0018)	
3 Clerical, Sales	0.637***	0.025***	**0.045**	10 Finance and	0.721**	0.012**	**0.020**
	(0.210)	(0.0094)		Large Business	(0.31)	(0.0060)	
4 Blue Collar	0.530**	0.0127*	**0.027**	11 Police, Army,	-0.617	-0.0027	**0.0041**
	(0.265)	(0.0076)		Security, Fire	(0.844)	(0.0031)	
5 Personal	0.578	0.013*	**0.025**	12 Professional,	-0.760	-0.0018	**0.0098**
Services	(0.268)	(0.0073)		Doctor, Lawyer	(0.488)	(0.0046)	
6 Teaching	0.371	0.018	**0.063**	13 Semi-skill Tech,	1.199***	0.011***	**0.0074**
	(0.181)	(0.011)		Call Centers	(0.387)	(0.0037)	
7 Government	0.671	0.0022	**0.0045**	14 Nursing, Health,	0.955***	0.014***	**0.015**
	(0.641)	(0.0025)		Hospital	(0.33)	(0.0051)	

Estimations include fixed effects at the household level. Marinal effects, dy/dx, are from corresponding multinomial logit estimations; control variables are gender, age, age^2, birthorder, and oldest child. Number of observations=4,956. Pseudo $R^2 = 0.0165$, Chi-squared $p < 0.0001$.

Table 2.9: Impact on Adult Wealth

	Dwelling Quality						
Variables	(1) Indoor toilet	(2) Electricity in home	(3) Improved walls	(4) Improved roof	(5) Improved floor	(6) Simple Dwelling Index	(7) Anderson Dwelling Index
OLS, Household FE:							
Sponsored	0.009	0.029***	0.025**	0.004	0.019**	0.082***	0.034*
	(0.006)	(0.007)	(0.012)	(0.006)	(0.009)	(0.017)	(0.020)
Observations	9,477	9,490	7,863	8,554	8,614	10,004	10,004
R-squared	0.006	0.008	0.011	0.004	0.012	0.013	0.009
GMM-IV, Household FE:							
Sponsored	-0.017	0.041	0.006	-0.001	0.070	0.232**	0.192*
	(0.028)	(0.035)	(0.049)	(0.030)	(0.049)	(0.093)	(0.106)
Observations	9,477	9,490	7,863	8,554	8,614	10,004	10,004

	Consumer Durables						
Variables	(8) Mobile phone	(9) Owns bike	(10) Owns motorcycle	(11) Owns car	(12) Owns land	(13) Simple Consumer Index	(14) Anderson Consumer Index
OLS, Household FE:							
Sponsored	0.054***	0.015	0.010	-0.000	0.003	0.089***	0.003
	(0.012)	(0.010)	(0.009)	(0.007)	(0.009)	(0.025)	(0.029)
Observations	9,884	9,856	9,906	9,880	9,444	10,004	10,004
R-squared	0.047	0.044	0.036	0.023	0.047	0.097	0.047
GMM-IV, Household FE:							
Sponsored	0.183***	0.004	0.019	0.000	-0.006	-0.004	0.024
	(0.055)	(0.052)	(0.040)	(0.036)	(0.046)	(0.128)	(0.145)
Observations	9,883	9,856	9,906	9,880	9,444	10,004	10,004
R-squared	0.085	0.060	0.106	0.029	0.063	0.149	0.077

Regressions include controls for age, gender, and sibling order. OLS and IV estimations incorporate household fixed effects. First-stage F-test for instrumental variable estimation yields $F = 95.82$ ($p < 0.001$). Clustered standard errors at the household level in parentheses.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$

Note: Whenever one of the components is missing for the particular house characteristic, we replace it with the mean for the index variables. This was done due to concern that creating an index only for individuals that had none of the constituent categories missing would lead to a non-representative sample. The "improved walls" variable, which is mildly significant, has the most missing observations. Dropping that variable slightly lowers the estimated impact of the index variables.

Table 2.10: Impact on Adult Wealth, Formerly Sponsored Men

	Dwelling Quality						
Variables	(1) Indoor toilet	(2) Electricity in home	(3) Improved walls	(4) Improved roof	(5) Improved floor	(6) Simple Dwelling Index	(7) Anderson Dwelling Index
OLS, Household FE:							
Sponsored	0.021**	0.030**	0.019	0.014*	0.045***	0.114***	0.081***
	(0.009)	(0.013)	(0.018)	(0.008)	(0.015)	(0.029)	(0.031)
Observations	4,829	4,833	4,007	4,377	4,413	5,048	5,048
R-squared	0.009	0.008	0.011	0.010	0.017	0.024	0.013
GMM-IV, Household FE:							
Sponsored	-0.043	0.079	-0.076	-0.066	0.111	0.132	0.047
	(0.039)	(0.053)	(0.063)	(0.048)	(0.084)	(0.133)	(0.149)
Observations	4,829	4,833	4,007	4,377	4,413	5,048	5,048

	Consumer Durables						
Variables	(8) Mobile phone	(9) Owns bike	(10) Owns motorcycle	(11) Owns car	(12) Owns land	(13) Simple Consumer Index	(14) Anderson Consumer Index
OLS, Household FE:							
Sponsored	0.073***	0.015	0.012	-0.000	-0.010	0.094**	-0.018
	(0.021)	(0.020)	(0.014)	(0.012)	(0.015)	(0.043)	(0.050)
Observations	4,986	4,971	4,993	4,979	4,806	5,048	5,048
R-squared	0.039	0.008	0.018	0.025	0.066	0.067	0.038
GMM-IV, Household FE:							
Sponsored	-0.048	-0.005	-0.029	0.031	-0.068	-0.154	-0.007
	(0.098)	(0.090)	(0.062)	(0.060)	(0.069)	(0.202)	(0.231)
Observations	9,883	9,856	9,906	9,880	9,444	10,004	10,004

Regression include controls for age, gender, and sibling order. OLS and IV estimations incorporate household fixed effects. First-stage F-test for instrumental variable estimation yields $F = 76.48$ ($p < 0.001$). Clustered standard errors at the household level in parentheses.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$

Table 2.11: Impact on Adult Wealth, Formerly Sponsored Women

	Dwelling Quality						
Variables	(1) Indoor toilet	(2) Electricity in home	(3) Improved walls	(4) Improved roof	(5) Improved floor	(6) Simple Dwelling Index	(7) Anderson Dwelling Index
OLS, Household FE:							
Sponsored	0.008	0.029**	0.026	-0.001	0.009	0.052*	-0.002
	(0.009)	(0.013)	(0.019)	(0.009)	(0.016)	(0.030)	(0.030)
Observations	4,648	4,657	3,856	4,177	4,201	4,956	4,956
R-squared	0.010	0.007	0.010	0.014	0.014	0.006	0.010
GMM-IV, Household FE:							
Sponsored	-0.020	-0.007	0.034	0.011	0.004	0.122	0.166
	(0.025)	(0.028)	(0.049)	(0.028)	(0.047)	(0.091)	(0.107)
Observations	4,648	4,657	3,856	4,177	4,201	4,956	4,956

	Consumer Durables						
Variables	(8) Mobile phone	(9) Owns bike	(10) Owns motorcycle	(11) Owns car	(12) Owns land	(13) Simple Consumer Index	(14) Anderson Consumer Index
OLS, Household FE:							
Sponsored	0.058***	0.012	0.018	0.003	0.006	0.107***	0.029
	(0.018)	(0.012)	(0.012)	(0.010)	(0.013)	(0.035)	(0.041)
Observations	4,898	4,885	4,913	4,901	4,638	4,956	4,956
R-squared	0.040	0.005	0.005	0.025	0.057	0.053	0.027
GMM-IV, Household FE:							
Sponsored	0.103	0.002	0.050	-0.011	0.021	0.143	0.031
	(0.065)	(0.044)	(0.042)	(0.035)	(0.048)	(0.129)	(0.145)
Observations	4,898	4,885	4,913	4,901	4,638	4,956	4,956

Regressions include controls for age, gender, and sibling order. OLS and IV estimations incorporate household fixed effects. First-stage F-test for instrumental variable estimation yields $F = 78.88$ ($p < 0.001$). Clustered standard errors at the household level in parentheses.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$

Table 2.12: Impacts on Home Residence

Multinomial Logit Estimations
(Base Category: Living in Parent's Home)

	All Individuals		Men		Women	
Variables	(1) Live in Rented Home	(2) Live in Owned Home	(3) Live in Rented Home	(4) Live in Owned Home	(5) Live in Rented Home	(6) Live in Owned Home
Multinomial Logit Coeffs:						
Sponsored	0.362***	0.161*	0.382***	0.089	0.318***	0.211*
	(0.085)	(0.088)	(0.118)	(0.128)	(0.121)	(0.121)
Marginal Effects:						
Sponsored	0.048***	0.00	0.057***	-0.011	0.035**	0.011
	(0.013)	(0.130)	(0.019)	(0.019)	(0.017)	(0.018)
Observations	8,365		4,288		4,077	
Pseudo R-squared	0.170		0.166		0.181	

Multinomial logit estimations include controls for age, gender, oldest child, sibling order, number of siblings, mother's education, father's education, and country fixed effects. Data on residence not obtained in Uganda. Baseline values among untreated: 46.8% living in parent's home, 23.8% living in rented home, 29.4% living in an owned home.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$

Chapter 3

INTERNAL MIGRATION IN BRAZIL: THE ROLE OF JOB MOBILITY AND GENDER

Laine Rutledge

Abstract

I analyze the return to migration and how it varies with firm change and gender. A common assumption in the migration literature is that individuals change firms upon changing locations, whereas I am able to distinguish between geographical movements occurring within and between firms. I then separate these effects by gender. I utilize a matched employer-employee dataset of Brazilian workers in the formal sector from 2003 to 2013. Due to the matched aspect of the data, I am able to include individual and firm fixed effects in the wage regressions. Women experience positive returns to moving within a firm when only individual fixed effects are included. Once unobserved firm characteristics are controlled for through the use of firm fixed effects, the positive returns to migration for women are eliminated. In either case, women do not see a statistically significant return to moving and changing firms. The addition of firm fixed effects does not greatly impact the return to moving for men. This indicates that any promotion and geographical movement that women experience is not reflected in earnings.

3.1 Introduction

Mobility is often seen by economists as an important means of escaping poverty. In the context of developing countries, this is typically framed as individuals moving from rural areas to urban areas, or from the agricultural sector to the industrial sector, where higher wages are typically paid. Much research has shown this to be a common migration pattern in developing countries, particularly in the informal sector, with individuals seeing large increases in income (or consumption) after a move. This research, in contrast, examines migration in the formal labor market and how returns to moving vary with firm changes. Previous research has been unable to link individuals to firms when analyzing migration outcomes, ignoring any effects unobserved firm characteristics have on wage determination. I use matched employer-employee data in the Brazilian formal sector in order to include firm fixed effects in addition to individual fixed effects in wage regressions that include an indicator for having moved. The inclusion of firm fixed effects in addition to individual fixed effects allows for disentanglement of characteristics important to wage determination. Recent literature has shown this to be important in various labor market contexts, but has yet to be applied to the study of migration. For example, Abowd et al. (1999) found a positive relationship between high-wage firms and increased productivity and profit, while high-wage workers were only found to have a positive relationship with increased productivity when including firm fixed effects. Additionally, Abowd et al. (2006) examine the return to seniority in firms and find a null result, however there exists significant heterogeneity between low and high starting-wage firms.

Because the migration process can be quite different for women, earlier research on migration has focused mainly on determinants and outcomes for men. Traditionally, women have been viewed as tied movers, only moving because their husbands move, rather than responding to their own economic opportunities. More recent literature has shown marital status and family size to be important determinants of migration for women, while outcomes are mixed. I first estimate returns to moving with and without firm changes on the entire sample. I then interact these different types of moves with gender in order to compare outcomes for men and women for similar types of moves.

Results for estimations across genders show the greatest return for individuals who relocate geographically but stay with the origin firm. A positive, but smaller, return is experienced by individuals who change both locations and firms, while local job changers experience a negative return. Adding various firm and geographic controls decreases all returns, with the smallest coming from the regressions that include firm fixed effects. When matched employer-employee data is not available, any estimated effect is the average across individuals who change firms and those who do not. These results show that estimates on moving may be overstated when not accounting for firm change. Also, the results vary significantly between genders. While including firm fixed effects decreases the effect of moving in the pooled regression, the results are quite different once the estimates are separated by gender. Adding firm fixed effects decreases the effect of moving but remaining with the firm from 6.7% to 6.0% for men, and slightly increases the effect of moving and changing firms from 3.8% to 3.9%. However, when firm fixed effects are added to the regressions for women, the effect of moving without changing firms goes from 5.2% to 2.8% but statistically insignificant. Women moving and changing firms see statistically insignificant returns regardless of whether firm fixed effects are included. This result lends important insight into the migration process for women because unobserved firm characteristics seem to play a large role in their returns to moving, but, up to this point, have not been included in migration estimations. I provide evidence that firm characteristics are indeed important in determining returns to migration, and play a more significant role in outcomes for women. Moreover, I show that migration coupled with a firm change is not a positive return strategy for women.

The remainder of the paper is structured as follows: Section 2 discusses relevant literature, section 3 presents the empirical methodology and discusses the data, section 4 reviews the results, and section 5 concludes.

3.2 Relevant Literature

Microeconomic theory has historically viewed migration as a response to economic opportunity. This idea was formally introduced by Sjaastad (1962), where migration decisions are influenced by employment opportunities and/or higher wages in various destination locations. As outlined by Roy (1951), the sample of individuals choosing to migrate is not random. Rather, those individuals self-select into the migration process on both observable and unobservable characteristics. Borjas (1987) formalizes the self-selection mechanism in the Roy (1951) model and combines it with the income-maximizing human capital method of Sjaastad (1962). The main contribution of Borjas (1987) is to show that, while regional income gaps indeed impact migration decisions, so do differences in returns to skill. The model predicts that less educated individuals will migrate to regions where returns to education are less dispersed and more educated individuals will migrate to regions where returns to education are more dispersed.

Various methods attempting to account for the self-selection of individuals into migration have been carried out over the years. For example, Gries et al. (2011) uses a Heckman selection model with family size in the selection equation and finds evidence of self-selection in individuals moving from the southeast region of Brazil to the northeast, but no evidence of self-selection in those moving in the opposite direction. Beegle et al. (2011) uses consumption data from Tanzania to study the impact of migration on living conditions using factors like baseline location, rainfall shocks, and individual status within the household as instruments in a two-stage least squares (2SLS) estimator to solve for the selection problem. Stillman et al. (2015) and Bryan et al. (2014) use experimental data to overcome the problem of not observing wage outcomes of movers had they not moved and self-selection, respectively. Bryan et al. (2014) use cash incentives to individuals in Bangladesh to encourage out-migration in the lean, pre-harvest season. They find that 22% of households are induced to send a migrant, consumption at the origin increases, and treated households are more likely to migrate again in later years. Stillman et al. (2015) utilizes an international migration lottery to show that international migration brings substantial increases in objective welfare measures. The data in this analysis is nonexperimental, but does have the advantage of including firm fixed effects in addition to individual fixed effects. This is advantageous because it allows me to control for firm-specific unobservables affecting the wage, as well as individual unobservables that do not vary as they move from firm to firm or location to location. The estimated return to migration, then, is net of individual ability, genetic health endowments, firm compensation policy, or firm management quality. These are all examples of unobserved characteristics at the individual or firm level that do not vary with time and affect wages.

Another important component of selection into migration is the presence of networks. Munshi (2003) shows that migration is greatly influenced by home networks and access to social networks in destination locations, and Kilic et al. (2009) offer evidence of community out-migration rates as an important determinant of migration. Both papers analyze international migration, where these effects are likely more pronounced than in the internal migration context. Similarly, McKenzie and Rapoport (2010) reveal the selection pattern of Mexico - U.S. migration to vary across communities depending on their networks.

In contrast to the international migration setting, where firm change is assumed (and almost guaranteed), internal migration studies are able to more closely examine the relationship between firm and geographic mobility. Bartel (1979) uses U.S. household data to show that the relationship between wages and migration is significantly dependent on the fact that job separations accompany migration. She argues the decision to change jobs can be a driving force in the decision to move, with movers who transfer seeing the largest returns and movers who are laid off seeing the smallest returns. There is also significant evidence that job tenure reduces the probability of job separation due to the positive relationship between tenure and job-specific training (Bartel and Borjas, 1977; Jovanovic, 1979). Yankow (2003) and Ham et al. (2005) use information on job separation to form different control groups, arguing that individuals who change jobs locally form a more appropriate counterfactual to out-migration since within-firm wage growth and between-firm wage growth are different.

Research on internal migration in developing countries typically shows large returns to migration, either in terms of income or consumption (Beegle et al., 2011; Bryan et al., 2014), and is framed as rural to urban or agricultural to industrial sector migration (Falaris, 1979). These results are not surprising given that many of the households surveyed are living at subsistence levels initially. In the case of Brazil, there were large internal migration streams between the 1960s and 1990s originating from the poorer northeast region going to the richer southeast region (Wagner and Ward, 1980; Yap, 1976). However, Gries, Kraft and Pieck (2011) show that this pattern of migration reversed itself in the 2000s, with the southeast region seeing an inflow of over 390,000 migrants during the 1990s and an outflow of 150,000 in 2006. Interestingly, this reversal has not been accompanied by large changes in income characteristics of the regions. In fact, regional inequality remains large in Brazil (Azzoni, 2001; Azzoni and Servo, 2002; Laurini, Andrade and Valls Pereira, 2005). Employment opportunities and higher wages offered by exporting firms in Brazil have been shown to be important determinants of migration flows (da Silva Freguglia et al., 2014; Hering and Paillacar, 2015). Because Brazil's main export goods are intensive in low-skilled labor, low-educated individuals have a greater response in changes to firms access to foreign markets , resulting in migration flows that are inverse to those assumed in the informal sector (Hering and Paillacar, 2015). However, these measures of migration streams include individuals working in both the formal and informal sectors. da Silva Freguglia et al. (2014) uses data on the formal labor market in Brazil, and does not find this pattern of migration. Likewise, this pattern is not evident in the RAIS data.

Much of the previous literature on migration, both internal and international, is focused on determinants and outcomes for male migrants. Research examining outcomes for interregional migration of women in developed countries shows wage penalties upon moving, likely due to the fact that women are often considered to be tied movers (moving with their husbands) (Grant and Vanderkamp, 1980; Lichter, 1983; Maxwell, 1988). Gries et al. (2011) concludes that women are paid less than men upon moving, using household data from Brazil. Fields (1979) analyzes migration propensities and patterns for men and women in Colombia, showing them to be highly similar. Oliveira (2016) specifically analyzes the determinants and outcomes for women migrating in Brazil and finds that an additional child reduces the migration rate by 6 percent and, consistent with previous literature, fertility is shown to have a negative impact on wages. However, 2SLS estimates show migrant women experience a 26% increase in hourly wages upon moving. I find positive returns to migration for women when only including individual fixed effects, but once firm fixed effects are included in the estimation these returns disappear.

The advantage to using household data when studying migration is that individuals working in both the formal and informal sectors are observed. Additionally, family characteristics such as marital status and number of children are observed and can be used to correct for self-selection into migration. However, despite only tracking individuals in the formal sector, linked employer-employee data offers other advantages. The panel does not suffer from attrition issues when individuals move, and most importantly, firm fixed effects can be included in the estimation. This allows for the disentanglement of the effect of firm-level decisions from the effects of choices made by workers, both important in wage determination. Methods for the inclusion of these effects have been introduced by Abowd et al. (2002), Guimarães and Portugal (2010), and Correia (2016). However, to my knowledge, previous literature has not included individual and firm fixed effects in wage regressions examining the effect of migration. If unobserved firm characteristics (e.g. unique technologies, compensation policies, quality of management) are uncorrelated with wages then pooled regressions will provide consistent estimates of the impact of moving. This is unlikely and the exclusion of these important variables may overstate the return to moving.

3.3 Empirical Framework and Data

3.3.1 Data

The data for this research was provided by the Ministry of Labor and Employment (MTE) in Brazil and the Brazilian Institute of Geography and Statistics (IBGE). Every formally registered firm in Brazil is required by law to respond to the Relação Anual de Informações Sociais (RAIS), or Annual Social Information Survey, administered by the MTE. Firm administrators provide the information in RAIS at the establishment level. This is typically the owner or manager in smaller firms, and a dedicated human resources employee in larger firms. Rosenzweig (2003) points out that most panel data used for the study of migration is likely to suffer from attrition, resulting in the loss of some of the most important households. This problem is avoided using the RAIS data, since full responses from firms are legally mandated. However, I am only able to measure returns to migration conditional on working in the formal sector before and after moving. If workers are moving geographically and from the formal sector to the informal sector due to displacement, estimates of the return to moving could be upwardly biased. However, Carneiro and Henley (2001) provide evidence that the informal sector in Brazil may be a desirable alternative to employment in the formal sector, rather than a result of cyclical displacement.

Firms provide information on all individuals formally employed in each year. This includes any individual who was employed for part of the year, if not all. Formal employment in Brazil is defined for workers who have a registered identification number either with the Programa de Integração Social (PIS), or Social Integration Program, for employees of private enterprises or a registered Programa de Formação de Patrimônio do Servidor Público (PASEP), or Equity Formation Program for Civil Servants. These identification numbers follow a worker for life, regardless of whether the worker is in PIS or PASEP. This allows researchers to track individuals across jobs. Formal employment is defined for employers where the employer contributes to a bank account administered by Banco do Brasil, for PASEP workers, or Caixa Econômica Federal for PIS workers, and the employment is recorded by a number of different employment contracts. The most common employment contract is known as the Consolidação das Leis de Trabalho (CLT), or Labor Law Consolidation. There are other contract types including internships, independent contractors, and non-owner executives, all of which are represented in the RAIS data.

In Brazil, formal employment entitles workers to constitutionally mandated benefits. These include vacation days (30 per year), maternity leave, disability and unemployment insurance, meal assistance, and the Abono Salarial, which is also known as the thirteenth salary or Christmas bonus. The Abono Salarial is equal to one month's pay if the worker was employed for the entire calendar year, and is prorated for those that were employed for less time. The MTE administers the thirteenth salary, and it is for this reason that they require firms to respond to the RAIS survey. Firms who do not respond open themselves up to lawsuits from employees not receiving their thirteenth salary, so coverage is universal.

Demographic information on employees includes age, race, gender, nationality, and education. Education is provided categorically, rather than by years of schooling. There are also variables describing whether a worker is classified as disabled and, if so, what type of disability. Job characteristics for employees include the date of hire, the type of hire (first employment, additional employment, temporary employment, etc.), tenure (in months), the date and cause of separation from the firm (if applicable), occupation, industry, and earnings. Occupation is according to the CBO 2002 classification system provided by the IBGE. I aggregate this code to the one-digit level resulting in the occupations listed in Table 3.1. Industry is according to CNAE 1995 or CNAE 2.0 classification systems, which I also aggregate to the 16 industry categories listed in Table 3.2. There are two production and manufacturing industries listing, the first represents all production in non-petrochemical products and the second consists of petrochemical production. Earnings are reported nominally as well as in multiples of the minimum wage. They are calculated by taking worker's total earnings for the year and dividing by 12. If the worker was not employed for the full calendar year, the division is adjusted appropriately so that the resulting value represents one month's pay. Hourly earnings are calculated by dividing the reported monthly wage by 4.17 to obtain a weekly wage, then dividing by hours worked per week. The obtained hourly wage is then deflated using the IPCA, a Brazilian consumer inflation index, provided by the IBGE.

3.3.2 Sample Construction

The data is originally provided as text files at the state-year level, for years 2003 through 2013. When combined, these files form a dataset of about 200 GB and over 650 million observations. Due to limited computing resources, I use an approximately 0.1 percent sample of the data for analysis. In order to create this sample, I aggregate the data to the yearly level, then randomly select workers from each year's data file. I create a list of all worker identifications randomly selected from each year and merge these back with the original yearly data sets, in order to create a panel of workers across years. This yields a panel data set with approximately 1.1 million observations. I then limit the data to workers with a positive wage, nonnegative tenure, and those aged 15 to 65. I also limit the data to native Brazilians, which results in dropping less than 2 percent of the sample. Until this point, the unit of observation is the job-year, but I am interested in the worker-year data so I limit the sample to one job per worker per year. However, since I am interested in wage changes, I calculate each individual's effective wage per year (weighted average of the hourly wage rate) prior to limiting the sample. I then eliminate jobs the worker is no longer employed at within a firm each year. The point of this step is to properly track job changes. Imagine the data look like the following:

Worker	Firm	Year	Currently Employed
1	1	2006	Yes
1	1	2007	No
1	2	2007	Yes

I eliminate the second observation because the individual is no longer employed at firm 1, but has gained employment with firm 2. Their employment at firm 1 is already recorded in 2006, and I want to capture the move from firm 1 to firm 2 in 2007. Random selection is used amongst the remaining multiple observations per worker per year to limit the data to a worker-year panel. These remaining observations do not vary in tenure, hours worked, or wage. Since I have calculated effective hourly wage, a single observation from these workers will suffice. Removing these observations results in a final sample with 76,769 individuals and 671,056 worker-year observations.

3.3.3 Migrant Definition

Municipality is reported for each worker, which I then aggregate to the state level. Municipalities can be quite small, so to more accurately define migrants I use state residence. For instance, I do not want to include commuters in my sample of migrants. An individual is then determined to be a migrant if state of residence in time $t{+}1$ is different from that in time t. An individual is coded as changing firms when the firm ID in time $t{+}1$ is different from the firm ID in t. Table 3.3 describes different types of moves and firm changes for individuals. Approximately 16 percent of the population migrates at some point between 2003 and 2013, 95 percent of whom change firms when they move and 83 percent stay within the firm upon migrating.[1] Of migrants, about 75 percent are male and 25 percent are female.

Table 3.4 shows summary statistics for the entire sample, migrants, and non-migrants

[1]These sum to over 100 percent because individuals move multiple times, sometimes moving within firms and sometimes between firms.

separately. The average effective hourly wage in the sample is R$12.69 (2013 Reals), which at the time would have been about 5.00 USD. Non-migrants average R$0.21 less than migrants. Men are more likely to be migrants than women, 59% of non-migrants are male versus 75% of migrants. Migrants are also younger than non-migrants, by about 3 years, and have far less tenure. Tenure amongst non-migrants is about 65 months and only 24 months for migrants. This is consistent with previous literature.

Tables 3.1, 3.2, and 3.5 show occupation, industry, and education frequencies amongst the full sample, non-migrants, and migrants. A larger proportion of non-migrants are professionals, 12.59% versus 8.45%, while 11.94% of the overall sample are professionals. There is a larger percentage of migrants that work in agriculture, 6.7% as compared to 4.26% of non-migrants. Of the entire sample, 4.64% work in agriculture. There are two different production and manufacturing categories. The first is production and manufacturing of non-petrochemicals and the second is petrochemical production and manufacturing. Approximately 22% of all individuals work in the production of non-petrochemicals, while 20.45% of non-migrants are employed in this occupation compared to 28.09% of migrants. The largest proportion of individuals are in the public administration industry, 19.74%. However, 21.95% of non-migrants are in public adminstration, while only 7.87% of migrants are in that industry. There is a much larger proportion of migrants in construction, 15.57% versus 4.4%, while 6.15% of the entire sample work in construction. Similarly, 19.16% of migrants are in real estate as compared to 11.37% of non-migrants, with 12.59% of the entire sample working in real estate. The most frequently observed education level is high school completion. Forty-two percent of the sample completed high school, with 42% and 43% of non-migrants and migrants completing high school, respectively. Only 4.59% of the sample has no or incomplete elementary education, while 4.32% of non-migrants and 6.05% of migrants having no or incomplete elementary education.

3.3.4 Empirical Methodology

To estimate the impact of migration on wages in a manner that more carefully controls for individual and firm characteristics, I employ five main regression specifications: ordinary least squares (OLS) adding individual fixed effects, ordinary least squares adding individual and firm fixed effects, both of these with an interaction term for migration and firm change status, and the previous four interacted with gender. The purpose of interacting these with gender is to explore any differences in outcomes between men and women, since the migration process may be quite different for women (e.g. women are moving with their husbands). There may also be differences in labor market conditions for women.

Individual Fixed Effects

Supplementing OLS with individual fixed effects controls for unobserved, time-invariant characteristics of the individual that may be important to wage determination, such as genetic health and ability endowments or risk aversion. The individual fixed effects equation for individual i in time t is given by:

$$y_{it} = \theta_i + x_{it}\beta + M_{it}\tau + \varepsilon_{it} \tag{3.1}$$

where y_{it} is the natural logarithm of the effective hourly wage rate observed for individual $i = 1, ..., N$ at time $t = n_{i1}, ..., n_{iT_i}$, θ_i captures the individual fixed effect, x_{it} is a vector of time-varying characteristics measured for individual i at time t including age, age-squared, tenure, disability status, education, occupation, industry, firm size, and state dummies, M_{it} is an indicator for whether individual i has moved at time t, and ε_{it} is the statistical residual. The indices $n_{i1}, ..., n_{iT_i}$ indicate the period corresponding to the first observation on individual i through the last observation on that individual, respectively, and T_i is the total number of years individual i appears in the sample. The coefficient of interest is τ, which gives the overall return to migration regardless of whether the individual changes firms when moving.

Much of the previous literature either implicitly or explicitly assumes that workers change firms upon migration. However this assumption does not hold in this context, confounding the impacts of migration and firm change on wage determination. While migrants do change jobs more frequently than non-migrants, 39% of migrants change jobs between 2003 and 2013 and 21% of non-migrants do so, a non-trivial number of migrants stay with the firm when changing location (4.68%). Interacting the indicator for changing firms with that for having moved allows me to separate the two effects. This is similar to the analysis in Böheim and Taylor (2007), however they do not have firm identifiers and instead use a subsample of individuals who report changing jobs as a counterfactual group. The estimate is as above, modified to include the main effect of firm change and the interaction term:

$$y_{it} = \theta_i + x_{it}\beta + M_{it}\tau + CF_{it}\phi + (M_{it} * CF_{it})\delta + \varepsilon_{it} \tag{3.2}$$

where the difference-in-difference coefficient on the interaction term, $\boldsymbol{\delta}$, reveals how the return to moving varies by firm change status, and y_{it}, θ_i, and x_{it} are as described above.

Because the migration process can differ substantially for men and women, I also interact gender with the indicators for moving and firm change. The main effect of gender is absorbed by the pure person effect since it is time-invariant, and does not need to be included explicitly in the estimation equation. The resulting equation is given by:

$$\begin{aligned} y_{it} = \theta_i + x_{it}\beta + M_{it}\tau + CF_{it}\phi + (M_{it} * F_i)\pi + (M_{it} * CF_{it})\delta + \\ (M_{it} * CF_{it} * F_i)\gamma + \varepsilon_{it} \end{aligned} \tag{3.3}$$

where F_i is an indicator for being female, π describes how the effect of moving varies with gender, γ reveals how the effect of moving varies with firm change status and gender, and all other elements are as previously described.

Individual and Firm Fixed Effects

In addition to controlling for individual effects that do not vary over time or as the individual moves from firm to firm, employer-employee linked data allows the inclusion of firm effects that control for time-invariant firm characteristics that also do not vary as the firm employs different individuals. The estimate including person and firm fixed effects is given by:

$$y_{it} = \theta_i + \psi_{\mathbf{J}(i,t)} + x_{it}\beta + M_{it}\tau + \varepsilon_{it} \tag{3.4}$$

where $\psi_{\mathbf{J}(i,t)}$ is the pure firm effect, the function $\mathbf{J}(i,t)$ indicates the firm $j = 1, ..., J$ of individual i at time t, and all other components are as defined previously. I will outline the individual and firm fixed effects estimation using the above equation in order to simplify notation, however, note that I also estimate the above equation including the interaction terms shown in equations 3.2 and 3.3. Equation 3.4 is what is known in the statistical literature as a two-factor analysis of covariance with two high-dimensional factors and a non-orthogonal design (Abowd et al., 2002). High-dimensional factors are defined as fixed effects with many categories that increase as sample size increases (Guimarães and Portugal, 2010). Putting equation 3.4 in matrix notation yields:

$$y = D\theta + F\psi + X\beta + M\tau + \varepsilon \tag{3.5}$$

where all vectors and matrices have row dimensionality equal to the total number of observations, except β which has row dimensionality equal to the number of time-varying controls, $N^* = \sum_i T_i$, D is the $N^* \times N$ design matrix for the individual effect, F is the $N^* \times J$ design matrix for the firm effect, X is the $N^* \times K$ stacked matrix of time-varying characteristics, and M is the $N^* \times 1$ vector of migration indicators. The full least squares solution to the estimation for equation 3.5 solves the following normal equations for all estimable effects:

$$\begin{bmatrix} X'X & X'D & X'F & X'M \\ D'X & D'D & D'F & D'M \\ F'X & F'D & F'F & F'M \\ M'X & M'D & M'F & M'M \end{bmatrix} \begin{bmatrix} \beta \\ \theta \\ \psi \\ \tau \end{bmatrix} = \begin{bmatrix} X'y \\ D'y \\ F'y \\ M'y \end{bmatrix} \tag{3.6}$$

In my estimation sample, the cross-product matrix on the left-hand side of equation 3.6 is too high-dimensional to use conventional algorithms implemented in Stata (or other statistical software packages). In fact, the traditional method of applying the within transformation to the fixed effect with more categories and including dummy variables for the fixed effect with less categories is not feasible with large datasets or when there is more than one high-dimensional fixed effect. In this case, the sample consists of 671,056 observations, which includes 76,769 individuals and 120,830 firms, so traditional computational methods are not possible. Abowd, Kramarz and Margolis (1999) discuss several different estimators that seek to maintain as much of the general structure of the estimation as computationally possible. A critical component necessary for the estimation and interpretation of the analysis is that there is inter-firm mobility. Independent of the computational approach taken, inter-firm mobility is necessary for the identification of the statistical model. Approximately 76% of individuals in my sample change firms at some point, and 93% of individuals are employed by firms that employ others who have switched firms at some date. Applying methods from graph theory to form groups of connected workers and firms, and then employing conventional analysis of covariance methods within connected groups allows for the identification of the individual and firm fixed effects (Abowd et al., 1999). Using an algorithm to construct G mutually-exclusive groups of connected workers and firms and the conjugate gradient method, Abowd et al. (2002) are able to estimate individual and firm effects. However, this process can be very time intensive for such a large data set.

3.3.5 Computational Expediency

Correia (2016) outlines a feasible estimator in the presence of multiple high-dimensional fixed effects that is computationally efficient. It begins by exploiting the insight in Guimarães and Portugal (2010) of applying the Frisch-Waugh-Lovell theorem, which implies that the least squares estimates $\hat{\boldsymbol{\beta}}$ can be recovered by first regressing each variable against all the fixed effects, then regressing the residuals of these variables. Following Correia (2016), let $\boldsymbol{P_D} = \boldsymbol{D}(\boldsymbol{D'D})^{-1}\boldsymbol{D'}$ be the projection matrix with respect to $\boldsymbol{D}$ and $\boldsymbol{M_D} = \boldsymbol{I} - \boldsymbol{P_D}$ the corresponding annihilator matrix. Then the vectors $\tilde{\boldsymbol{y}} = \boldsymbol{M_D y}$ and $\tilde{\boldsymbol{X}} = \boldsymbol{M_D X}$ are the residuals of $\boldsymbol{y}$ and $\boldsymbol{X}$ with respect to the fixed effects. The two step process then consists of:

1. $\hat{\boldsymbol{\beta}} = (\tilde{\boldsymbol{X}}'\tilde{\boldsymbol{X}})^{-1}\tilde{\boldsymbol{X}}'\tilde{\boldsymbol{y}}$

2. $\tilde{\boldsymbol{\varepsilon}} \equiv \boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}} + \boldsymbol{D}\hat{\boldsymbol{\alpha}} = \tilde{\boldsymbol{y}} - \tilde{\boldsymbol{X}}\hat{\boldsymbol{\beta}}$

In the presence of only one fixed effect, this simplifies to the traditional within-estimator, and $\boldsymbol{M_D}$ will simply subtract group means. The remaining step is to obtain the OLS residuals of a model of the form $\boldsymbol{y} = \boldsymbol{D\alpha} + \boldsymbol{\varepsilon}$ for $k+1$ variables. In terms of normal equations, it is equivalent to solving the linear system:

$$(\boldsymbol{D}'\boldsymbol{D})\hat{\boldsymbol{\alpha}} = (\boldsymbol{D}'\boldsymbol{y}) \tag{3.7}$$

Guimarães and Portugal (2010) implement a block version of Von Neumann (1949)'s and Halperin (1962)'s Method of Alternating Projections that demeans each variable across a fixed effect, obtains the residuals, and then repeats cyclically across all fixed effects until the residuals converge to the partialled-out variables. While it has guaranteed convergence, it can be quite slow despite acceleration with conjugate gradient. Correia (2016) utilizes the fact that the $\boldsymbol{D}'\boldsymbol{D}$ matrix belongs to a class of matrices that can be reduced to Laplacian matrices, which can be solved in "nearly-linear time", to construct a computationally efficient estimator.[2] This allows for the estimation of the variables of interest, while controlling for the high-dimensional firm and worker fixed effects, in a timely manner.

3.4 Empirical Results

First, I will examine the overall return to migration. Then I will move on to results analyzing how the return to moving varies with firm separation. Although previous work has used firm

[2]See Correia (2016) for details on Laplacian matrices and graph theory as it applies to high-dimensional fixed effects estimation issues.

separation as a way of refining the control group, it has not been interacted with moving in this way before. Doing so lends some insight to the overall effect of moving. Lastly, I will discuss the results that include a triple interaction term of moving, firm separation, and gender. The outcomes for women migrants have received less attention in the literature, and certainly not considered in combination with firm separation.

3.4.1 Return to Migration

Table 3.6 presents the impact of moving on logged effective hourly wage from moving, not controlling for firm change. Column (1) shows the effect of moving to be a 5.5% increase in wage without including any firm or geographic controls. Columns (2) and (3) further decompose the effect of moving by adding in firm size dummies and geographic characteristics, respectively. Adding firm covariates decreases the return to moving to 4.6%, and including geographic characteristics further decreases the return to 4.0%. Individual fixed effects controls for any unobserved individual heterogeneity, but adding a pure firm effect allows me to disentangle the effects of choices made by workers and firm-level decisions. Column (4) shows that the addition of firm fixed effects decreases the return to moving to 3%. This is a greater decrease than seen in the first three columns when only controlling for observed firm characteristics. This suggests that unobserved firm heterogeneity is an important contributor to wage determination in the migration context, even without controlling for firm change.

3.4.2 Return to Migration and Firm Change

To further disentangle the effect of moving on wages, I interact the indicator for having moved with a that for changing firms. This allows for firm heterogeneity in measuring the return to moving by allowing the return to vary depending on whether the individual is also changing firms. Table 3.7 presents results from this interaction, with average marginal effects below the regression results, and columns (1), (2), (3) and (4) decompose the result as in Table 3.6. The return to moving without changing firms, which can be thought of as a transfer or promotion, is substantially higher than the return to moving and changing firms. Those who move but remain with the firm see an increase in wage of 7.7% to 4.4%, while those who switch firms when moving see an increase of 4.6% to 2.7%. This is not a surprising result, as individuals who transfer within a firm likely have job-specific training that still applies in the new location. Note also that the coefficient on changing firms locally is negative across estimations, with little variation depending on whether firm or geographic controls are included, even when firm fixed effects are added in column(4). The average marginal effects reported in column (4) exhibit larger decreases in the return to moving, both within and between firms, than seen in the previous three columns. When firm fixed effects are included a return of 4.4% results for those who remain with the firm and 2.7% for those who change firms. This is similar to the pattern seen when adding firm fixed effects to the estimation of a simple dummy for moving on wages. Again, the unobserved firm characteristics absorb some of the positive return to moving. Figure 3.1 plots the average marginal effects and confidence intervals of moving by firm change status for all four estimations. The average marginal effect of moving and changing firms steadily declines as additional controls are added, with the smallest return resulting from the inclusion of both individual and firm fixed effects. The pattern is different for those who do not change firms upon moving, however. Steady decreases are seen as firm and geographic controls are added, with a larger jump happening when individual and firm fixed effects are included.

3.4.3 Return to Migration, Firm Change, and Gender

The migration process can be quite different for men and women for various reasons. Women may not decide to move in response to economic opportunity, but instead move because their husband is moving. The returns to human capital also vary between men and women.

Because of this, I estimate the previous equations but include an interaction of moving, changing firms, and a gender dummy. Because occupation and industry effects also differ by gender, I estimate the equations with the full set of interactions. Table 3.8 presents the results of gender and moving, without controlling for firm change, with average marginal effects at the bottom of the table. Column(1) shows a 4.4% increase in wages for men who move under individual fixed effects. This decreases to 4.3% when firm fixed effects are added in column(2). Women see a 3% increase in wages upon moving when only including individual fixed effects, but this falls to 2.1% and is not statistically signifcant when controlling for firm fixed effects. Notice that including firm fixed effects only decreased the return to males by 0.1%, but eliminated the return for females. The large decrease seen when including firm fixed effects but without controlling for gender appears to be driven by this outcome on returns to females.

Table 3.9 presents the results of the three-way interaction across two estimations, with average marginal effects reported below the regression results. The excluded group is those who not move or change firms, conditional on gender. Column (1) reports results of the individual fixed effects estimation on the three-way interaction. Males see a 6.7% increase when remaining with the firm and a 3.8% increase when changing firms. Females see a 5.2% increase when remaining with the firm, but no significant return when changing firms, suggesting that moving without a job in place is not a positive return strategy for women. This could be due the fact that women often move only because their husband is moving. Comparing these results to column (2), where both fixed effects are included, males who move within a firm experience a 6% increase in wage (only a 0.7% decrease from the result of the individual fixed effects estimation). Women who remain with the firm, however, see no statistically significant increase in wage, whereas the return was 5.2% when only including individual fixed effects. Additionally the return to changing firms locally is negative for both genders, but less so for women. Figure 3.2 plots the average marginal effects and confidence intervals of the interaction term. The estimations for men are more precise, with the return decreasing slightly when firm fixed effects are included. For both men and women, the return increases slightly when adding firm fixed effects, in the case of changing firms when moving. However, the result is not statistically significant for women, as mentioned. While including firm fixed effects leads to predictable results on the interaction between moving and firm change, the results from separating the outcomes of males and females are quite different when accounting for firm effects.

3.5 Conclusion

Despite a vast literature on the outcomes of migration, little attention has been paid to how this outcome varies with firm separation. Additionally, firm fixed effects have not previously been applied in the context of migration. By incorporating this information into the estimations of migration on wages, this research provides some insight on the role of firm characteristics on wage determination in the migration context. These estimations do not account for the self-selection component of migration, and, as such, the return may be overestimated. However, the main insight of this research is the importance of firm characteristics on wage determination of men versus women in a migration context, and the inclusion of these characteristics fully absorbs the positive returns accruing to women when a simple individual fixed effects estimator is used.

The result that men experience a positive return to migration and women do not when facing the same unobservable firm characteristics indicates that women are either not being promoted when moving within a firm, are not compensated equally for these types of promotions, firms have discriminatory compensation policies, or women themselves are initiating transfers as opposed to being promoted or transferred by the firm. Implementing legislation for gender-neutral practices, incentivizing promotion of women, or prohibiting gender-based discrimination in firms could correct for the lack of returns experienced by women and contribute to economic growth. Additionally, previous literature has shown that women do not respond to economic opportunity in other locations as often as men do, underlining the importance of positive returns upon seeking out these opportunities.

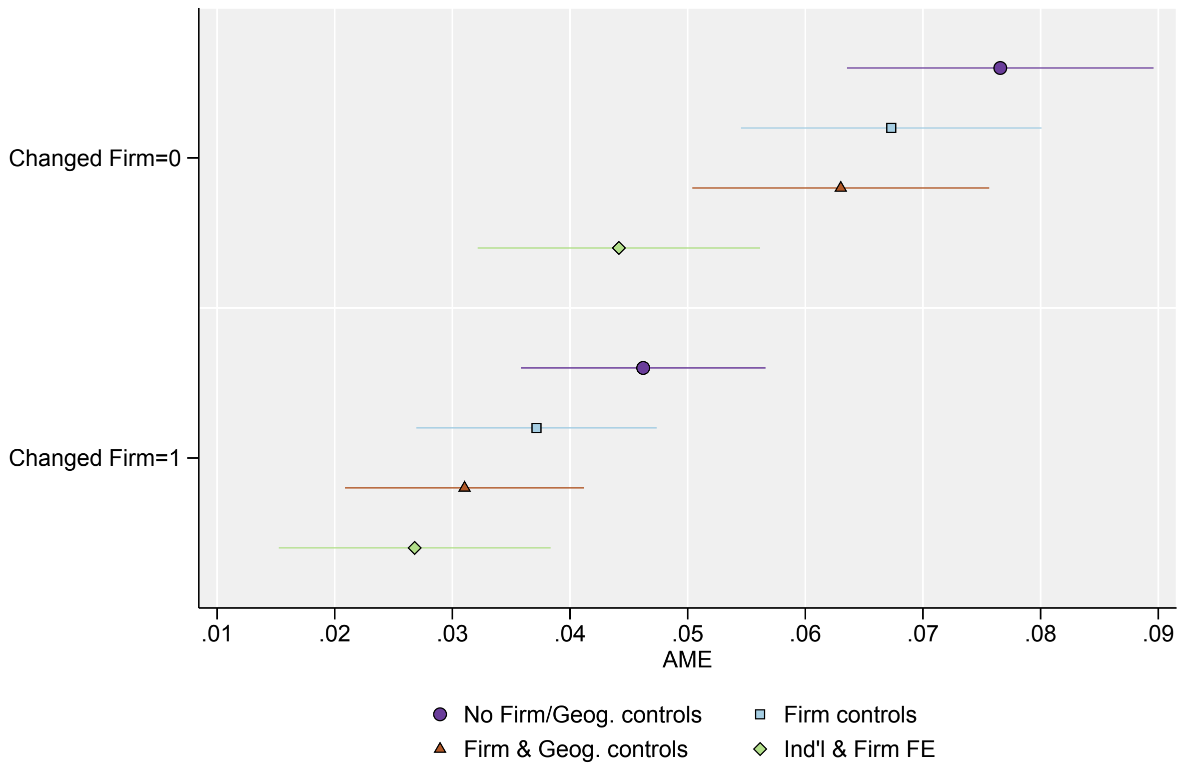

Figure 3.1: Average Marginal Effect of Moving by Firm Change Status

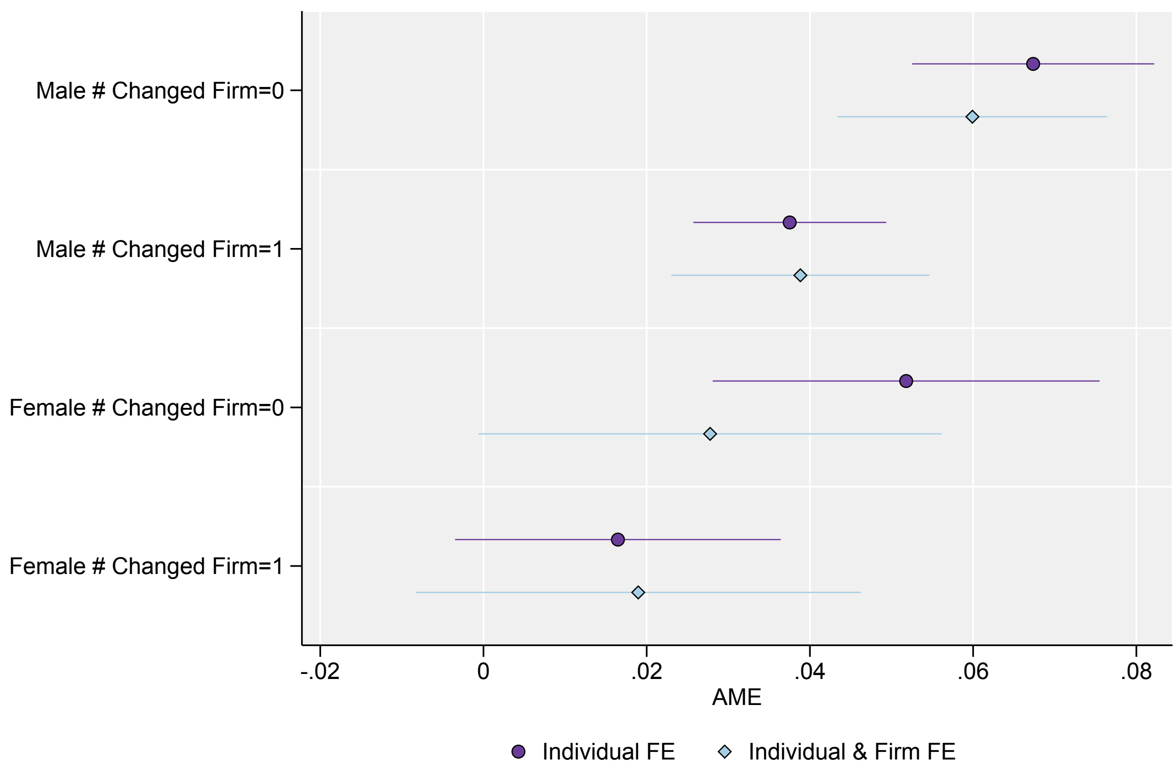

Figure 3.2: Average Marginal Effect of Moving by Firm Change Status and Gender

Table 3.1: Occupation Frequencies

	All	Non-migrants	Migrants
Public administration and management	4.39	4.40	4.31
Professionals, artists, and scientists	11.94	12.59	8.45
Mid-level technicians	11.55	11.86	9.90
Administrative workers	18.25	18.57	16.54
Service workers and vendors	21.04	21.37	19.26
Agriculture workers, fishermen, forestry workers	4.64	4.26	6.70
Production and manufacturing, I	21.65	20.45	28.09
Production and manufacturing, II	3.57	3.68	3.02
Repair and maintenance workers	2.97	2.82	3.74
Total	100.00	100.00	100.00
Observations	664956	560498	104458

Table 3.2: Industry Frequencies

	All	Non-migrants	Migrants
Agriculture and Fishing	4.11	3.83	5.60
Mining	0.49	0.43	0.81
Production/manufacturing	17.93	18.34	15.74
Utilities	0.53	0.55	0.39
Construction	6.15	4.40	15.57
Trade and repair	17.48	17.77	15.94
Food, lodging, and hospitality	3.00	3.05	2.75
Transportation, storage, and communication	5.48	5.26	6.69
Financial and intermediary services	1.88	1.75	2.59
Real estate, renting, and services	12.59	11.37	19.16
Public admin, defense, and public security	19.74	21.95	7.87
Education	3.50	3.75	2.12
Health and social services	3.82	4.16	2.00
Other social and personal services	3.27	3.37	2.74
Domestic services	0.01	0.01	0.01
International and extra-territorial organizations	0.01	0.01	0.01
Total	100.00	100.00	100.00
Observations	671054	565514	105540

Table 3.3: Migration and Job Mobility

	Ind'ls	Person-years	% Total	% Migrants	% Non-migrants
All	76,769	671,056	100	-	-
Non-Migrants	64,319	565,516	84.27	0	100
Migrants	12,450	105,540	15.73	100	0
Male Migrants	9,333	79,176	11.80	75.02	0
Female Migrants	3,117	26,364	3.93	24.98	0
Changed firm (only)	46,260	405,417	60.41	-	71.69
Moved and changed firm	11,874	100,096	14.92	94.84	0
Males	8,887	74,949	11.17	71.01	0
Females	2,987	25,147	3.75	23.83	0
Moved and stayed with firm	9,857	87,830	13.09	83.22	0
Males	7,295	65,209	9.72	61.79	0
Females	2,562	22,621	3.37	21.43	0

Table 3.4: Summary Statistics

	All	Non-migrants	Migrants	Difference
Avg. Hourly Wage	12.69	12.66	12.87	-0.21**
	(24.01)	(24.07)	(23.68)	(0.08)
Male	0.62	0.59	0.75	-0.16***
	(0.49)	(0.49)	(0.43)	(0.00)
Age	35.76	36.25	33.16	3.09***
	(10.45)	(10.58)	(9.30)	(0.03)
White	0.65	0.67	0.58	0.09***
	(0.48)	(0.47)	(0.49)	(0.00)
Metro Area	0.63	0.63	0.63	0.00**
	(0.48)	(0.48)	(0.48)	(0.00)
Tenure	58.83	65.26	24.40	40.86***
	(79.94)	(83.45)	(43.63)	(0.26)
Observations	671056	565516	105540	

Table 3.5: Education Frequencies

	All	Non-migrants	Migrants
None or incomplete elementary	4.59	4.32	6.05
Elementary	13.99	13.89	14.58
Middle/primary	21.18	21.17	21.23
High School	42.10	41.94	42.91
Post-secondary	18.14	18.68	15.24
Total	100.00	100.00	100.00
Observations	670830	565314	105516

Table 3.6: Impact of Moving on Wage, Not Controlling for Firm Change

	(1) Ind'l FE	(2) Ind'l FE	(3) Ind'l FE	(4) Ind'l & Firm FE
Moved	0.055***	0.046***	0.040***	0.030***
	(0.005)	(0.005)	(0.005)	(0.006)
Tenure	0.001***	0.001***	0.001***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)
Disabled	-0.019***	-0.015**	-0.014**	-0.008
	(0.005)	(0.005)	(0.005)	(0.005)
Age	0.060***	0.056***	0.056***	0.047***
	(0.001)	(0.001)	(0.001)	(0.002)
Age-squared	-0.001***	-0.001***	-0.001***	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)
Elementary	0.008	0.014***	0.013**	0.006
	(0.004)	(0.004)	(0.004)	(0.005)
Middle/primary	-0.005	0.010*	0.008	0.002
	(0.004)	(0.004)	(0.004)	(0.005)
High school	0.011*	0.023***	0.021***	-0.000
	(0.005)	(0.005)	(0.005)	(0.006)
Post-secondary	0.167***	0.169***	0.166***	0.106***
	(0.007)	(0.006)	(0.006)	(0.007)
Metro Area			0.045***	0.001
			(0.004)	(0.006)
Occupation dummies	Yes	Yes	Yes	Yes
Firm size dummies	No	Yes	Yes	No
Industry dummies	No	Yes	Yes	No
State dummies	No	No	Yes	Yes
Observations	637073	637071	637071	622199
R^2	0.334	0.350	0.353	0.288

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3.7: Impact of Moving on Wage, Controlling for Firm Change

	(1) Ind'l FE	(2) Ind'l FE	(3) Ind'l FE	(4) Ind'l & Firm FE
Moved=1	0.077***	0.067***	0.063***	0.044***
	(0.007)	(0.007)	(0.006)	(0.006)
Changed Firm=1	-0.020***	-0.019***	-0.019***	-0.019***
	(0.001)	(0.001)	(0.001)	(0.001)
Moved=1 × Changed Firm=1	-0.030***	-0.030***	-0.032***	-0.017***
	(0.005)	(0.005)	(0.005)	(0.003)
Metro Area			0.046***	0.001
			(0.004)	(0.006)
Occupation dummies	Yes	Yes	Yes	Yes
Firm size dummies	No	Yes	Yes	No
Industry dummies	No	Yes	Yes	No
State dummies	No	No	Yes	Yes
Average Marginal Effects				
Moved=1 × Changed Firm=0	0.077***	0.067***	0.063***	0.044***
	(0.007)	(0.007)	(0.006)	(0.006)
Moved=1 × Changed Firm=1	0.046***	0.037***	0.031***	0.027***
	(0.005)	(0.005)	(0.005)	(0.006)
Observations	637073	637071	637071	622199
R^2	0.334	0.351	0.353	0.289

Standard errors in parentheses

Tenure, disability status, age, age-squared, and education controls included

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3.8: Female and Male Migration, Not Controlling for Firm Change

	(1) Ind'l FE	(2) Ind'l & Firm FE
Moved=1	0.044***	0.043***
	(0.006)	(0.008)
Female × Moved=1	-0.014	-0.022
	(0.012)	(0.016)
Occupation dummies	Yes	Yes
Firm size dummies	Yes	No
Industry dummies	Yes	No
State dummies	Yes	Yes
Average Marginal Effects		
Moved=1 × Male	0.044***	0.043***
	(0.006)	(0.008)
Moved=1 × Female	0.030**	0.021
	(0.010)	(0.014)
Observations	637071	596591
R^2	0.355	0.291

Standard errors in parentheses

Tenure, disability status, age, age-squared, and education controls included

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3.9: Female and Male Migration, Controlling for Firm Change

	(1) Ind'l FE	(2) Ind'l & Firm FE
Moved=1	0.067***	0.060***
	(0.008)	(0.008)
Changed Firm=1	-0.025***	-0.024***
	(0.002)	(0.002)
Moved=1 × Changed Firm=1	-0.030***	-0.021***
	(0.006)	(0.005)
Female × Moved=1	-0.016	-0.032
	(0.014)	(0.017)
Female × Changed Firm=1	0.013***	0.013***
	(0.003)	(0.003)
Female × Moved=1 × Changed Firm=1	-0.005	0.012
	(0.011)	(0.009)
Occupation dummies	Yes	Yes
Firm size dummies	Yes	No
Industry dummies	Yes	No
State dummies	Yes	Yes
Average Marginal Effects		
Moved=1 × Male × Changed Firm=0	0.067***	0.060***
	(0.008)	(0.008)
Moved=1 × Male × Changed Firm=1	0.038***	0.039***
	(0.006)	(0.008)
Moved=1 × Female × Changed Firm=0	0.052***	0.028
	(0.012)	(0.014)
Moved=1 × Female × Changed Firm=1	0.016	0.019
	(0.010)	(0.014)
Moved=0 × Male × Changed Firm=1	-0.025***	-0.024***
	(0.002)	(0.002)
Moved=0 × Female × Changed Firm=1	-0.012***	-0.011***
	(0.002)	(0.002)
Observations	637071	596591
R^2	0.355	0.292

Standard errors in parentheses

Tenure, disability status, age, age-squared, and education controls included

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

BIBLIOGRAPHY

Aaronson D, Mazumder B. 2011. The Impact of Rosenwald Schools on Black Achievement. *Journal of Political Economy* **119**: 821–888.

Abowd JM, Creecy RH, Kramarz F. 2002. Computing Person and Firm Effects Using Linked Longitudinal Employer-Employee Data. Technical report, Center for Economic Studies, US Census Bureau.

Abowd JM, Kramarz F, Margolis DN. 1999. High Wage Workers and High Wage Firms. *Econometrica* **67**: 251–333.

Abowd JM, Kramarz F, Roux S. 2006. Wages, Mobility and Firm Performance: Advantages and Insights from Using Matched Worker-Firm Data. *The Economic Journal* **116**: F245–F285.

Anderson ML. 2008. Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects. *Journal of the American statistical Association* **103**.

Azzoni CR. 2001. Economic Growth and Regional Income Inequality in Brazil. *The annals of regional science* **35**: 133–152.

Azzoni CR, Servo L. 2002. Education, Cost of Living and Regional Wage Inequality in Brazil. *Papers in regional science* **81**: 157–175.

Bank W. 2009. Conditional Cash Transfers: Reducing Present and Future Poverty. Technical report, Policy Research Report. The World Bank, Washington, D.C.

Barrera-Osorio F, Bertrand M, Linden LL, Perez-Calle F. 2008. Conditional Cash Transfers in Education; Design Features, Peer and Sibling Effects: Evidence from a Randomized Experiment in Colombia. Working Paper 13890, NBER.

Bartel AP. 1979. The Migration Decision: What Role Does Job Mobility Play? *The American Economic Review* : 775–786.

Bartel AP, Borjas GJ. 1977. Middle-Age Job Mobility: Its Determinants and Consequences. Working Paper 161, National Bureau of Economic Research.
URL http://www.nber.org/papers/w0161

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

Behrman JR, Parker SW, Todd P. 2007. Do School Subsidy Programs Generate Lasting Benefits? A Five-Year Follow-Up of Oportunidades Participants. Working paper, University of Pennsylvania.

Behrman JR, Sengupta P, Todd P. 2005. Progressing through PROGRESA: An Impact Assessment of a School Subsidy Experiment in Rural Mexico. *Economic Development and Cultural Change* **54**: 237–275.

Bernard T, Dercon S, Taffesse AS. 2011. Beyond Fatalism: An Empirical Exploration of Self-Efficacy and Aspirations Failure in Ethiopia. Working Paper 2011-03, Center for the Study of African Economies.

Bertrand M, Karlan D, Mullainathan S, Shafir E, Zinman J. 2010. What's Advertising Content Worth? Evidence from a Consumer Credit Marketing Field Experiment. *The Quarterly Journal of Economics* **125**: 263–306.

Blattman C, Fiala N, Martinez S. 2014. Generating Skilled Self-Employment in Developing Countries: Experimental Evidence from Uganda. *Quarterly Journal of Economics* **129**: 697–752.

Bobonis GJ, Finan F. 2009. Neighborhood Peer Effects in Secondary School Enrollment Decisions. *The Review of Economics and Statistics* **91**: 695–716.

Böheim R, Taylor MP. 2007. From the Dark End of the Street to the Bright Side of the Road? The Wage Returns to Migration in Britain. *Labour Economics* **14**: 99–117.

Borjas GJ. 1987. Self-Selection and the Earnings of Immigrants. *National Bureau of Economic Research Cambridge, Mass., USA* .

Bryan G, Chowdhury S, Mobarak AM. 2014. Underinvestment in a Profitable Technology: The Case of Seasonal Migration in Bangladesh. *Econometrica* **82**: 1671–1748.

Carneiro F, Henley A. 2001. Modelling Formal vs. Informal Employment and Earnings: Micro-Econometric Evidence for Brazil. Working Paper 2001 - 15, U of Wales at Aberystwyth Management and Business.

Correia S. 2016. A Feasible Estimator for Linear Models with Multi-Way Fixed Effects. Working paper.

Cristia J, Ibarrarán P, Cueto S, Santiago A, Severín E. 2012. Technology and Child Development: Evidence from the One Laptop per Child Program. Working Paper IDB-WP-304, IDB.

da Silva Freguglia R, Gonçalves E, da Silva ER. 2014. Composition and Determinants of the Skilled Out-Migration in the Brazilian Formal Labor Market: A Panel Data Analysis from 1995 to 2006. *Economia* **15**: 100–117.

Dalton PS, Ghosal S, Mani A. 2010. Poverty and Aspirations Failure: A Theoretical Framework. Working paper.

Drèze J, Kingdon G. 2001. School Participation in Rural India. *Review of Development Economics* **5**: 1–24.

Duflo E. 2001. Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment. *The American Economic Review* **91**: 795–813. ISSN 00028282.
URL http://www.jstor.org/stable/2677813

Duflo E, Kremer M, Robinson J. 2011. Nudging Farmers to Use Fertilizer: Theory and Experimental Evidence from Kenya. *The American Economic Review* **101**: 2350–2390. ISSN 00028282.
URL http://www.jstor.org/stable/23045645

Dupas P. 2010. Short-Run Subsidies and Long-Run Adoption of New Health Products: Evidence from a Field Experiment. Working Paper 16298, NBER.

Evans D, Kremer M, Ngatia M. 2008. The Impact of Distributing Uniforms on Children's Education in Kenya. Working paper, World Bank.

Falaris EM. 1979. The Determinants of Internal Migration in Peru: An Economic Analysis. *Economic Development and Cultural Change* **27**: 327–341.

Fields GS. 1979. Lifetime Migration in Colombia: Tests of the Expected Income Hypothesis. *Population and Development Review* : 247–265.

Furth G. 2002. *The Secret World of Drawings: A Jungian Approach to Health through Art.* Inner City Books.

Glewwe P, Ilias N, Kremer M. 2010. Teacher Incentives. *American Economic Journal: Applied Economics* **2**: 205–227. ISSN 19457782, 19457790.
URL http://www.jstor.org/stable/25760225

Glewwe P, Miguel E. 2008. The Impact of Child Health and Nutrition on Education in Less Developed Countries. *Handbook of Development Economics, Volume 4* **North Holland**.

Glewwe P, Ross PH, Wydick B. 2014. Developing Hope: The Impact of International Child Sponsorship on Self-Esteem and Aspirations. Working paper, University of San Francisco.

Glewwe P, Wydick B. 2012. Child Sponsorship and Child Psychology: Evidence from Children's Drawings in Indonesia. Working paper, University of San Francisco.

Grant EK, Vanderkamp J. 1980. The Effects of Migration on Income: a Micro Study with Canadian Data 1965-71. *Canadian Journal of Economics* : 381–406.

Gries T, Kraft M, Pieck C. 2011. Interregional Migration, Self-Selection and the Returns to Education in Brazil. *The Annals of Regional Science* **46**: 707–732.

Guimarães P, Portugal P. 2010. A Simple Feasible Procedure to Fit Models with High-Dimensional Fixed Effects. *Stata Journal* **10**: 628.

Halperin I. 1962. The Product of Projection Operators. *Acta Sci. Math.(Szeged)* **23**: 96–99.

Ham JC, Reagan PB, Li X. 2005. Propensity Score Matching, a Distance-Based Measure of Migration, and the Wage Growth of Young Men. Working paper, IEPR.

Handa S, Peterman A. 2007. Child Health and School Enrollment: A Replication. *The Journal of Human Resources* **42**: 863–880. ISSN 0022166X.
URL http://www.jstor.org/stable/40057332

Heckman JJ. 1979. Sample Selection Bias as a Specification Error. *Econometrica* **47**: 153–161.

Hering L, Paillacar R. 2015. Does Access to Foreign Markets Shape Internal Migration? Evidence from Brazil. *The World Bank Economic Review* : lhv028.

Jovanovic B. 1979. Job Matching and the Theory of Turnover. *The Journal of Political Economy* : 972–990.

Kennedy J. 2009. The Not-for-Profit Surge. *Christianity Today* **53**: 22–27.

Kilic T, Crletto G, Davis B, Zezza A. 2009. Investing Back Home: Return Migration and Business Ownership in Albania. *Economics of Transition* **17**: 587–623.

Klepsch M, Logie L. 1982. *Children Draw and Tell: An Introduction to the Projective Uses of Children's Human Figure Drawings.* New York: Brunner-Routledge.

Koppitz E. 1968. *Psychological Evaluation of Children's Human Figure Drawings.* New York: Grune & Stratton.

Kremer M, Miguel E, Thornton R. 2009. Incentives to Learn. *The Review of Economics and Statistics* **91**: 437–456.

Kremer M, Moulin S, Namunyu R. 2003. Decentralization: A Cautionary Tale. Working Paper 10, Poverty Action Lab.

Kremer M, Vermeersch C. 2004. School Meals, Educational Attainment, and School Competition: Evidence From a Randomized Evaluation. Working Paper WPS3523, World Bank.

Laurini M, Andrade E, Valls Pereira PL. 2005. Income Convergence Clubs for Brazilian Municipalities: a Non-Parametric Analysis. *Applied Economics* **37**: 2099–2118.

Lichter DT. 1983. Socioeconomic Returns to Migration Among Married Women. *Social Forces* **62**: 487–503.

Maxwell NL. 1988. Economic Returns to Migration: Marital Status and Gender Differences. *Social Science Quarterly* **69**: 108.

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

Miguel E, Kremer M. 2004. Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities. *Econometrica* **72**: 159–217. ISSN 00129682, 14680262.
URL http://www.jstor.org/stable/3598853

Mullainathan S. 2006. Development Economics through the Lens of Psychology. Working paper, Harvard University.

Munshi K. 2003. Networks in the Modern Economy: Mexican Migrants in the United States Labor Market. *Quarterly Journal of Economics* **118**: 913–945.

Office of Management and Budget. 2012. Fiscal Year 2012 Budget of the U.S. Government. Technical report.

Oliveira J. 2016. Fertility, Migration, and Maternal Wages: Evidence from Brazil. *Journal of Human Capital* **10**: 377–398.

Ozier O. 2011. Exploiting Externalities to Estimate the Long-Term Effects of Early Childhood Deworming. Working paper, World Bank.

Pitt MM, Rosenzweig MR, Gibbons DM. 1993. The Determinants and Consequences of the Placement of Government Programs in Indonesia. *The World Bank Economic Review* **7**:

319–348. ISSN 02586770, 1564698X.
URL http://www.jstor.org/stable/3989824

Rawlins R, Pimkina S, Barrett CB, Pedersen S, Wydick B. 2014. Got Milk? The Impact of Heifer International's Livestock Donation Programs in Rwanda on nutritional outcomes. *Food Policy* **44**: 202–213.

Ray D. 2006. *Understanding Poverty*, chapter Aspirations, Poverty, and Economic Change. New York: Oxford University Press.

Rosenzweig M. 2003. Payoffs from Panels in Low-Income Countries: Economic Development and Economic Mobility. *American Economic Review* **93**: 112–117.

Ross P. 2010. An Investigation of Reference Point Shifts from a Child Sponsorship Program in Bolivia. Working paper, University of San Francisco.

Ross P, Wydick B. 2011. The Impact of Child Sponsorship on Self-Esteem, Life-Expectations, and Reference Points: Evidence from Kenya. Working paper, University of San Francisco.

Roy AD. 1951. Some Thoughts on the Distribution of Earnings. *Oxford economic papers* **3**: 135–146.

Schultz TP. 2004. School Subsidies for the Poor: Evaluating the Mexican Progresa Poverty Program. *Journal of Development Economics* **74**: 199–250.

Sjaastad LA. 1962. The Costs and Returns of Human Migration. *Journal of Political Economy* **70**: 80–93.

Stillman S, Gibson J, McKenzie D, Rohorua H. 2015. Miserable Migrants? Natural Experiment Evidence on International Migration and Objective and Subjective Well-Being. *World Development* **65**: 79–93.

Von Neumann J. 1949. On Rings of Operators. Reduction Theory. *Annals of Mathematics* **50**: 401–85.

Wagner F, Ward JO. 1980. Urbanization and Migration in Brazil. *American Journal of Economics and Sociology* : 249–259.

Wooldridge JM. 2010. *Econometric Analysis of Cross Section and Panel Data*. MIT press.

Wydick B, Glewwe P, Rutledge L. 2013a. Does Child Sponsorship Work? Evidence from a Six-Country Study on Adult Life Outcomes. Working paper, University of San Francisco.

Wydick B, Glewwe P, Rutledge L. 2013b. Does International Child Sponsorship Work? A Six-Country Study of Impacts on Adult Life Outcomes. *Journal of Political Economy* **121**: 393–436.

Yankow JJ. 2003. Migration, Job Change, and Wage Growth: a New Perspective on the Pecuniary Return to Geographic Mobility. *Journal of Regional Science* **43**: 483–516.

Yap L. 1976. Internal Migration and Economic Development in Brazil. *The Quarterly Journal of Economics* : 119–137.