International Child Sponsorship Impact on the Intended Choice of Schooling: the Case of Rural Mexico *

Daniel Prudencio Phillip Ross

Abstract

We study the impact of a child sponsorship program on the intended choice of schooling of children in the states of Oaxaca and Chiapas in Mexico. In order to account for the program's procedure to select the children, we estimate a binary Roy type model, with unobservables generated by a one-factor structure. We use both carefully scripted survey questions to elicit the subjective expected returns to schooling, and the results in Aakvik et al. (2005) to estimate the sponsorship effect on the revealed intended choice of acquiring a technical degree or university studies. We estimate the effect both on the population of eligible children for sponsorship and on the sponsored children, as well as the marginal sponsorship effect. The sponsorship effect is positively correlated with selection to the program, with an average sponsorship effect on the sponsored of 20 percentage points and a null effect on the average sponsorship effect. Nevertheless, the standard deviations are large, and both the average sponsorship effect on the sponsored and on the population are not statistically significant.

1 Introduction

Despite large returns to education, impoverished regions in Mexico still present very low levels of education. In two of the most socio-economically marginalized states, Oaxaca and Chiapas, the average years of education for adults older then 24 are 7.5 and 7.2, just above one year after primary school. Much thought and resources have been given to increase the level of education among the poor, and conditionally cash transfer programs focused on relieving external constrains, such as *Prospera* (previously known as *Oportunidades*), have shown positive effects in the level of education (Todd and Wolpin, 2006; Attanasio et al., 2012). It is well established that the presence of external constrains, such as the lack of income and poor health, is detrimental for the development of children, nevertheless, a growing body of research is also emphasizing that the presence of internal constraints, such as the lack of grit, perseverance, self-control, self-esteem and aspirations, are also

*Prudencio, Doctoral Student, Department of Economics Rice University, 6100 Main Street, Houston, TX 77005, email: danielprudencio@rice.edu. Ross, Doctoral Student: Department of Economics, Boston University, 270 Bay State Road, Boston, MA 02215, email: phross@bu.edu. We would like to thank Alistair Sim, Gloria Laux and other local Compassion staff for logistical help and support in carrying out our field research. We also appreciate support and helpful comments from Bruce Wydick, Flavio Cunha, and Rosella Calvi. All error are our own.

important inhibitors of the development of the children and youth (Cunha and Heckman, 2009; Heckman and Kautz, 2012; Heckman, Pinto and Savelyev, 2013; Dalton et al., 2016).

Interventions that relieve internal constraints may be important for the development of the child, especially at younger ages, as they can exacerbate the presence of external constraints, and because such constraints are more malleable at younger ages (Borghans et al, 2006). Internal constraints can take the form of psychological factors that restrain the development of the child, which can be reflected in the form of depression, hopelessness, a lack of hope or aspirations. Therefore, when analyzing programs that aim to enhance the development of children and youth, it is of interest to know if such programs have an effect on their internal constraints[1]. In this paper, we study the impact of a child international sponsorship program, called Compassion International, which was a holistic view of the development of the child. And we specifically analyzing its impact on the lack (or existence) of educational aspirations among the youth. We define the lack of aspirations as the revealed intention to not acquire higher education, defined as studying either a technical degree or at a university.

International child sponsorship programs serve as a means to transfer direct resources to impoverished children in developing countries. Additionally to the transfer of resources, some programs, as Compassion International (from now on CI), also focus on the relief of internal constraints through the development of socio-emotional skills of the children. In a recent study of CI, Glewwe et al., (2017) use children's self portraits to study the impact of sponsorship on the levels of hope, happiness and self-efficacy of the children, and find positive effects in the order of 0.66, 0.42 and 0.29 standard deviations of normalized indexes of the respective variables. Similarly, Wydick et al., (2013) found long term impacts of sponsorship by CI in school attainment (1.03 - 1.46 years), and on the probability of acquiring a white collar job (6.6 percent). In this line of research, we intend to study whether these long term impacts of child sponsorship may be due to the relief of both external constraints (through an in kind income transfer from CI), and internal constraints (through an increase in schooling aspirations).

Clearly, the relief of the external constraints may be the sole explanation of the program's effect on adult outcomes. Where we to find a null effect of the program on aspirations, we could rule out this channel as a main explanation for the adult outcomes observed. In this sense, we view this research as a necessary step, although not sufficient, to establish a link between aspirations and a higher level of education.

The sponsored individuals are selected through a clear selection procedure. In order to account for the selection used by CI to sponsor the children, we use a binary Roy type model with one factor structure to estimate the impact of CI on the intended choice of schooling. Furthermore, we use the results of Aakvik et al. (2005) for this specific model to estimate the mean and distributional treatment parameters for the average sponsorship effect on the population, and for the average sponsorship effect on the sponsored.

[1]For recent studies, see Heckman and Kautz (2012) and Heckman et al., (2013) who study the impact of the Perry Preschool Project. They authors study how the program may have been instrumental in the development of psychological factor of young children, and how this effects may have accrued in better outcomes during adulthood.

The present study also relates to the literature on subjective expectations. When studying the intended choice of schooling, we make use of expectations data to elicit the perceived returns to schooling, similar to Jensen (2010), Nguyen (2008), Attanasio and Kaufman (2009), Kaufman (2014) and Stinebrickner and Stinebrickner (2012, 2014). Also, the paper relates to a growing literature that uses the stated school choice rather than actual choices. Beliefs about educational future choices are of interest, since they tend to be one of the most important predictors of actual choices, above other standard determinants of schooling (Jacob and Wilder, 2011; Beaman et al., 2012). Recently, stated school choices have been used in the framework of discrete choice models to study the choice of the level of schooling (Delavende and Zafar, 2017), and the choice of a major (Arcidiacono, 2004; Arcidiacono et al., 2017, Zafar, 2013; Wiswall and Zafar, 2015).

The rest of the paper proceeds as follows, first, we introduce the institutional framework of CI, where we describe both how CI chooses the sponsored children and the benefits it provides. Second, we describe the field work and the data collected, specifically the module on subjective expectations. Next, we proceed to specify the model and the estimation of the sponsorship effect on the population, and on the sponsored individuals. We then discuss the results and conclude.

2 Institutional framework

Compassion International (CI) is the third largest child sponsorship program in the world, sponsoring around 1.3 million children throughout 26 countries. CI is a faith based, non profit organization that partners with local churches to carry its programming. The staff is local, and foreign workers are rare.

The benefits that the sponsored children receive vary little between countries. Mainly, they participate in structured programs after school that emphasize their spiritual, physical an socio-emotional development[2], and receive an in kind income transfer. In practice, this means that the students receive academic tutoring, school supplies and uniforms, classes that emphasize the spiritual and socio-emotional development of the child, health care in the form of general examinations by local nurses and doctors affiliated with CI, and a catastrophic health insurance. The usual sponsorship lasts until secondary school, but some receive support even during university through a leadership development program. In a large study of CI conducted over six different countries by Wydick et al.(2013), the average duration of sponsorship was 9.3 years, which accounts approximately for 4,000 hours of programmed activities for each sponsored child. CI has a large presence in Mexico, with 33,360 children participating in more than 185 child development centers.

In order to select the children for sponsorship, CI first selects a region within a country in which to operate, where they use the local poverty index and the human capital index to select the most marginalized regions. Then, within a region, they look for the poorest communities to open a center, and they partner with a church to conduct their programming. After CI opens a program in a community, the local staff recruits the neediest households from their community, and then the

[2]In Mexico, the program lasts between five to six hours per week, and it takes place during the weekend.

parents select the children that will be sponsored, provided that the children meet the following criteria:

- The household where the child lives is no further then 30 minutes walking distance from a CI center.
- The child is not receiving sponsorship through another organization.
- There must not be more then three sponsored children per household.
- The children must be 9 or younger to start receiving sponsorship. After age 10, the children are not eligible to start receiving sponsorship. Also, children close to 9 years old receive lower priority.

Preference is given to orphans, refugees and children that live with a widowed parent, or family member. Finally, although CI partners with churches to conduct its programming, CI sponsors children from both Christian and non-Christian families. Nevertheless, the family must allow the children to attend the program activities, which take place at the church.

3 Fieldwork and Data

The data was collected during the months of June to August of 2017, in the states of Oaxaca and Chiapas in the south of Mexico. The fieldwork was realized as part of a larger evaluation of CI, in which we surveyed eight communities, four with a CI project, and four without one. The CI sites where selected from rural or semi-rural settings, of which one was selected from the state of Oaxaca and three from Chiapas. A nearby community for each compassion site was selected based on similarity given observables. Importantly, the nearby sites had the same educational and health institutions as the CI sites, that is, each site had only one primary school, one middle-school, one high-school, and only one health post for basic care.

We surveyed the sponsored children at their households, along with their next oldest and next youngest sibling provided they where between 10 to 18 years old. Within the sites with a sponsorship program, we also surveyed households where non of its members where sponsored. We randomly selected the non-sponsored households, and surveyed all of the members between 10 to 18 years old of each selected household.

In both sites, with and without sponsorship, the sampling of the non-sponsored households proceeded as follows. After randomly choosing a starting point in the village, we selected every second household on the street for possible inclusion in the survey. When the end of the street or block was reached, the enumerators turned left and continued with every second household, then they turned right and proceeded in the same way. The household were briefly questioned to see if any member was between 10 to 18 years old, and we proceeded if there was a member that met the age criteria. Some surveyed individuals had siblings living in other parts of the village or in other communities. We tried to reach such individuals provided we had approval from their parents.

We gathered data on demographic characteristics, educational attainment, desired level of education, variables of self-esteem, optimism and aspirations based on Rosenberg (1965). For the level of income, since we did not surveyed the parents, we asked the children for household assets in order to build an assets index as an approximate for income, following Filmer and Pritchett (2001) (the authors build the asset index using the first principal component)[3]. Finally, in order to elicit the expected returns to education, we used data on subjective expectations, which we explain in more detail in the following section.

The initial sample consisted of 926 individuals aged 10 to 18. Given our use of expected income, and given that children in rural areas usually start to work after primary school when they are aged 12, we use only the data for the individuals aged 12 or older. Furthermore, the CI sites had in general less then 6 years since they started their operations, so I further restricted the sample to the individuals aged 12 to 15. Furthermore, I restricted the sample to the individuals that correctly interpreted the probability questions. A total of 271 individuals met the data criteria, of which 109 where sponsored and 162 where non-sponsored.

Some demographic characteristics for both the sponsored and the non-sponsored are presented in Table 1. Both groups where constrained to be between 12 to 15, nevertheless, differences arise in age by half a year, religion and the asset index, with the sponsored individuals showing a lower level. This differences are consistent with the selection criteria of CI, which is impacted by the age eligibility criteria, and the selection of the poorest families in a community. Apart from this variables, we do not find statistically significant differences in the rest of the demographic factors. For a more desegregated summary see Tables 5 and 6 in the Appendix, where we divide the data by the sponsorship status of the individual, separating the sponsored from the non-sponsored siblings of sponsored individuals, the non-sponsored individuals from non-sponsored households in a community with CI, and households in communities without the presence of CI. The observations remain the same. Also in the appendix in 7, we compare the individuals by community.

3.1 Subjective expectations data and estimation of the returns to schooling

Observed choices may be consistent with different specifications of preferences and expectations, so researchers commonly assume particular sorts of expectations. The measurement of these expectations could avoid this identification problem (Manski, 2004).

The use of subjective expectations is becoming more common, and their use has been applied for a wide array of topics, which include: income expectations (Dominitz and Manksi, 1997), social security expectations (Bernheim, 1998; Gustman and Steinmeier, 1999,2001), mutual fund investments (Dominitz and Manski, 2003,2004), probabilistic polling (Manski, 1990), and students expectations of the returns to schooling (Dominitz and Manski, 1996; Nguyen,2008; Attanasio and Kaufman,2013; Jensen, 2010; Kaufman, 2014, Stinebrickner and Stinebrickner, 2014; Zafar, 2011, 2013). For a review see Attanasio (2009), and for a review more focused on developing countries see Delavende et al. (2011).

[3]Some have used this index to study inequality in Mexico, see McKenzie (2005)

Table 1: Summary statistics: sponsored vs non-sponsored

	Mean all (std.dev.)	Sponsored (std.dev.)	Non-Sponsored (std.dev.)	t-test
Higher educ	0.765	0.679	0.731	-0.087
	(0.425)	(0.469)	(0.444)	(0.055)
University	0.660	0.550	0.616	-0.110*
	(0.475)	(0.500)	(0.487)	(0.061)
Age	13.623	13.046	13.391	-0.578***
	(1.098)	(1.013)	(1.100)	(0.129)
Male	0.481	0.413	0.454	-0.069
	(0.501)	(0.495)	(0.499)	(0.062)
Birth order	2.535	2.430	2.492	-0.105
	(1.905)	(1.632)	(1.798)	(0.235)
Num. siblings	4.000	4.055	4.022	0.055
	(1.939)	(2.085)	(1.996)	(0.284)
Protestant	0.352	0.798	0.531	0.446***
	(0.479)	(0.403)	(0.500)	(0.056)
Prospera	0.852	0.862	0.856	0.011
	(0.356)	(0.346)	(0.352)	(0.045)
Education father	6.727	7.085	6.873	0.358
	(3.320)	(3.193)	(3.267)	(0.439)
Education mother	6.704	6.101	6.461	-0.603
	(3.186)	(3.391)	(3.277)	(0.429)
Asset index	0.289	-0.217	0.085	-0.506***
	(1.356)	(1.294)	(1.352)	(0.181)
N	271	162	109	271

In this study, we elicit subjective expectations on the returns to schooling. We follow a similar procedure used in an evaluation of *Prospera* (previously known as *Oportunidades*), a well known conditional cash transfer program in Mexico. In 2002/03, a complementary program called *Jóvenes con Oportunidades* was introduced as part of *Oportunidades*, and subjective expectations data was collected in 2005. We use the same module here.

First, each individual was asked about the probability of working conditional on different schooling scenarios.

- Assume that you finish (*level of education*) and that this is your highest schooling degree. From 0 to 100, how certain are you that you will be working at the age of 25?

Then, the questions on subjective expectations of earnings conditional on a schooling level and on being employed, are the following:

- Assume that you finish (*level of education*) and that this is your highest schooling degree. Assume that you have a job at age 25.

  (a) What do you think is the maximum amount you can earn per month at that age?
  (b) What do you think is the minimum amount you can earn per month at that age?
  (c) From 0 to 100, what is the probability that your earnings at that age will be at least x, where $x = (\frac{max+min}{2})$?

This question was asked for five levels of education: primary school, middle school, high-school, technical studies and university. Following Atanassio and Kaufman (2013) who use the data on *Oportunidades*, we form the returns to education from the previous questions. Let Y^ℓ be the level of income with level of education ℓ. We are interested in the subjective distributions of future earnings $f(Y^\ell)$, where $(\ell = 1)$ for primary school,$(\ell = 2)$ for middle-school, $(\ell = 3)$ for high-school, $(\ell = 4)$ for technical studies, and $(\ell = 5)$ for college. Given that the survey asks for the maximum and the minimum of expected income for each level of education, we have the support of the distribution for each individual $[y^\ell_{min}, y^\ell_{max}]$, furthermore, we know $p = Pr(Y^\ell > y^\ell_{mid})$, for $y^\ell_{mid} = (y^\ell_{min} + y^\ell_{max})/2$. Thus, assuming a triangular distribution for $f(\cdot)$, we can estimate the expected value of the log of future earnings for each individual and level of education $(\ell = 1, 2, 3, 4, 5)$.

$$E[ln(Y^\ell)] = \int_{y_{min}}^{y_{max}} ln(y) f_{Y^\ell}(y) dy$$

Using these results, we can directly estimate the expected returns to schooling:

$$\rho^\ell = E[ln(Y^\ell)] - E[ln(Y^{\ell-1})], \text{ for } \ell = 2, 3, 4, 5.$$

3.2 Validity check of the subjective expectations data

Given that we have cross-section data, we can only compare the subjective expectation data with historical recordings, as opposed to comparing the expectations to realizations. Nevertheless, in this specific case, we can directly compare the subjective expectations to those from the survey on *Prospera* from 2005. As mentioned before, we are using the same question as in the *Prospera* survey, and furthermore, 85 per cent of our sample consists of households that receive *Prospera*.

Notice that differences between the subjective expectations data and historical realizations does not necessarily mean that the subjective expectations data are not valid. It is actually the fact that they may differ that we are interested in measuring the subjective expectations. The perceived expected returns to schooling matter for school choice, as it is the perceived return that may affect choices.

Next, we describe how the expectations data compares to census data of 2015 at prices of 2017. Table 2 displays the median of the data by gender and locality, where the census data has been further divided between regions with a population greater and lower than 50 thousand[4]. Various patterns are of interest. First, the median of the expectations data does not show unreasonable deviations, with higher income expectation in the state of Oaxaca than in Chiapas, consistent with the census recordings. Second, we do observe higher levels compared to those reported by Attanasio and Kaufman (2014), where in their sample of Prospera recepients, the individuals underestimated the level of income compared to census data, whereas in our sample the students overestimated it. This difference could be explained by the difference in age between the samples, theirs being a sample of individuals between 15 and 25, and ours of 12 to 15. Also, ours is a smaller sample. Third, even at the young ages of 12 to 15, a clear difference arise in expectations by gender, with the

[4]The population groups are formed given constraints in the census data set.

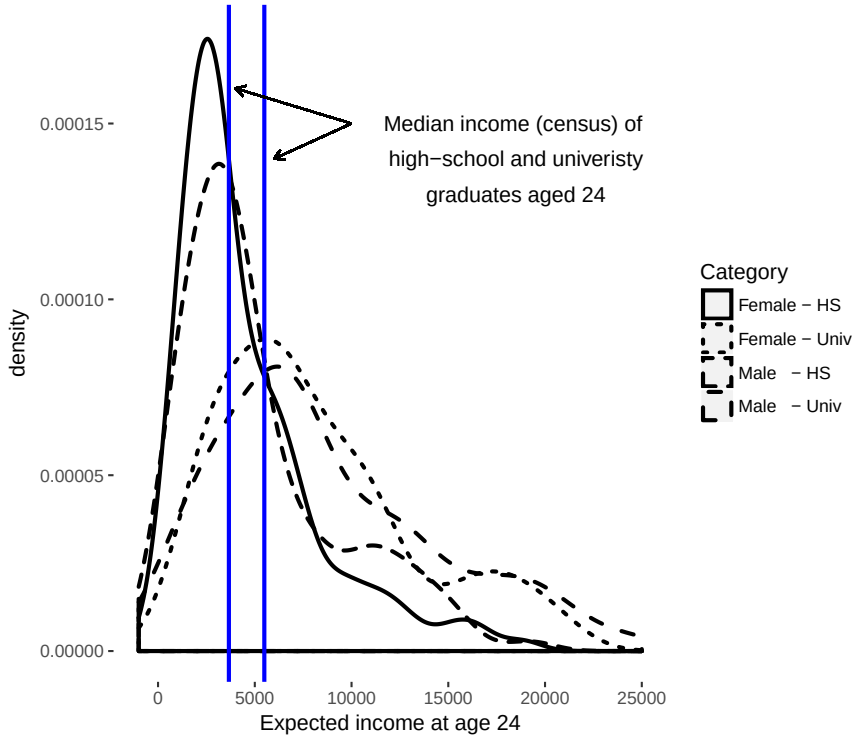
Table 2: Expected income vs census data at age 24 (median of data)

	High-school			Techical school			University		
	Male	Female	Total	Male	Female	Total	Male	Female	Total
Census									
Oaxaca-less50k	4271.7	3295.6	4119	5186.9	4271.7	4577.1	6407.6	6407.6	6407.6
	(4293.9)	(2577.1)	(3777.6)	(3296.6)	(2939.9)	(3123.7)	(3661.4)	(3058.3)	(3348.0)
Oaxaca-50k	4577.1	3867.5	4271.7	7933.1	4577.1	4577.1	4698.9	5339.6	5339.6
	(4359.6)	(1691.5)	(3549.4)	(5394)	(2008.4)	(3220)	(3052.2)	(3231.6)	(3151.8)
Chiapas-less50k	3203.8	2745.6	3203.8	3753.2	2867.9	3661.9	5339.6	4577.1	4752.3
	(2749.4)	(2300)	(2655)	(4208.7)	(2646.5)	(3656.1)	(3504.8)	(3042)	(3304.2)
Chiapas-50k	4271.7	3547.1	3844.5	6407.6	2563	5034.7	5339.6	5339.6	5339.6
	(2444.1)	(2397.4)	(2456.6)	(4876.8)	(3806.1)	(4372.2)	(27981.3)	(4658.4)	(20422.7)
Survey									
Oaxaca	5277.2	2796.5	3832.9	7701.2	3915.5	5493.7	11772.2	5939.7	7417.1
	(4772.6)	(14176.3)	(10567.5)	(7993.7)	(15996.3)	(12603.9)	(11293.2)	(19883.3)	(16134.1)
Chiapas	3823.9	3416.3	3587.3	5513.2	5277.2	5314.8	7429.7	7971.8	7574.7
	(5897.1)	(4347.2)	(5089.7)	(8761.4)	(5886.5)	(7319.8)	(14966.2)	(11228.3)	(13027)
AK(2014)-averages	3525.5	2509.3					6110.6	4806.8	

female individuals expecting future income to be around 50 percent less then males in Oaxaca, and 8 percent in Chiapas for future income with high-school and a technical education. Interestingly, female youth had higher expectations than males given university studies in the state of Chiapas, nevertheless, this relationship is reversed when looking at the means. A table comparing the means and logs rather then the medians can be found in the Appendix in Tables 8 and 9.

Finally, next in figure 1 we observe the sample distribution of believed future income by gender, where both the distribution for income given high-school and university studies has been graphed. Both distributions present fat tails to the right, but the modes are close to the census medians represented by the vertical lines, with the first line representing the median income of the high-school graduates at age 24, and the second line representing the median income of university graduates aged 24, in the states of Oaxaca and Chiapas.

Figure 1: Expected income distribution for high-school and university by gender



4 Model

To account for the selection of the sponsored individuals based on income, age and participation in activities within the church facilities, we use a discrete choice latent index framework following Aakvik et al. (2005), where we estimate a three equation model. The equations modeled are an outcome equation for the sponsored, an outcome equation for the non-sponsored, and a selection equation. Let S_i denote the sponsorship variable, where $S_i = 1$ if the individual i is sponsored, and $S_i = 0$ if non-sponsored. Let (Y_{1i}, Y_{0i}) be the two potential binary outcomes for individual i under sponsorship and non-sponsorship. Where $Y_{1i} = 1$ if individual i is sponsored and reveals that her intend choice of schooling is higher education, and $Y_{1i} = 0$ if she reveals an intended choice of schooling lower then higher education. Likewise, Y_{0i} is defined as Y_{1i} but for the non-sponsored. Then, the observed outcome variable is $Y_i = S_i Y_{1i} + (1 - S_i)Y_{0i}$.

We assume that a latent variable model generates S, such that

$$S_i^* = \gamma_0 + age\_6_i\gamma_1 + age\_7_i\gamma_2 + age\_8_i\gamma_3 + asset\_index_i\gamma_4 + religion_i\gamma_5 + site\_CI_i\gamma_6 + U_{Si},$$
$$S_i^* = Z_i\gamma + U_{Si},$$
$$S_i = \begin{cases} 1 \text{ if } S_i^* > 0, \\ 0 \text{ otherwise.} \end{cases}$$

where Z_i is a vector of observed covariates, $\gamma = (\gamma_0, ..., \gamma_6)$ the set of parameters of the selection equation, and U_{Si} is an unobserved random variable to the econometrician. As exclusion restric-

tions, similar to Wydick et al., (2013), We will use a set of dummy variables for the age of the individual when the CI project started at the community of the individual, here denoted as Age_p for $p = (6, 7, 8)$, where Age_p is a dummy variable for the individuals that were "p" years old when the program started in the community. The omitted category is Age_9 or older. We also consider whether the children attends a protestant church, $(religion_i)$, a dummy variable for the communities that have the sponsorship program $(site\_CI_i)$, and finally, an asset index using the first principal component of a set of household assets. Then $Z_i = (1, dage_i, asset\_index_i, religion_i, site\_CI_i)$. The outcome equation of the intended choice of schooling for the sponsored is:

$$\begin{aligned}
Y_{1i}^* &= \beta_0^1 + \rho_i \beta_2^1 + Dist_i \beta_3^1 + \tilde{X}_i \beta_4^1 + U_{1i}, \\
Y_{1i}^* &= X_i \beta_1 + U_{1i}, \\
Y_{1i} &= \begin{cases} 1 \text{ if } Y_{1i}^* > 0, \\ 0 \text{ otherwise.} \end{cases}
\end{aligned}$$

Likewise, the outcome equation of the intended choice of schooling for the non-sponsored is:

$$\begin{aligned}
Y_{0i}^* &= \beta_0^0 + \rho_i \beta_2^0 + Dist_i \beta_3^0 + \tilde{X}_i \beta_4^0 + U_{0i}, \\
Y_{0i}^* &= X_i \beta_0 + U_{0i}, \\
Y_{0i} &= \begin{cases} 1 \text{ if } Y_{0i}^* > 0, \\ 0 \text{ otherwise.} \end{cases}
\end{aligned}$$

Where $X_i = (1, \rho_i, Dist_i, \tilde{X}_i)$ are observed covariates, $\beta_0 = (\beta_0^0, ..., \beta_4^0)$, and $\beta_1 = (\beta_0^1, ..., \beta_4^1)$ are the coefficients of the outcome equations, and U_{1i} and U_{0i} are the unobserved random variables. Here ρ_i stand for the believed return to higher education, which was elicited directly as explained above. $Dist_i$ stands for the distance (in kilometers) to the closest public university, and $\tilde{X}_i$ is a set of other covariates such as gender, an asset index, and a dummy variable for whether one of the parents had higher then primary school education. Following Aakvik et al. (2005), we assume that the error terms U_{Si}, U_{1i}, U_{0i} follow a factor structure:

$$\begin{aligned}
U_{Si} &= -\theta_i + \varepsilon_{Si}, \\
U_{1i} &= -\alpha_1 \theta_i + \varepsilon_{1i}, \\
U_{0i} &= -\alpha_0 \theta_i + \varepsilon_{0i}.
\end{aligned}$$

We further assume that $(\theta_i, \varepsilon_{Si}, \varepsilon_{1i}, \varepsilon_{0i})' \sim N(0, I)$ for all i. Notice that the normalization of $Var(\theta) = Var(\varepsilon_1) = Var(\varepsilon_0) = 1$ is without loss of generality for the estimation of the sponsorship effects. To estimate the model, conditional on θ, the likelihood of the model is:

$$\begin{aligned}
L_i &= \prod_{i=1}^{N} Pr(S_i, Y_i | X_i, Z_i, \theta_i), \\
&= \prod_{i=1}^{N} Pr(Y_i | S_i, X_i, \theta_i) Pr(S_i | Z_i, \theta_i),
\end{aligned}$$

where it can be shown that

$$Pr(Y_i = 1|S_i = 1, X_i, \theta_i) = \Phi(X_i\beta_1 + \alpha_1\theta_i),$$
$$Pr(Y_i = 1|S_i = 0, X_i, \theta_i) = \Phi(X_i\beta_0 + \alpha_0\theta_i),$$
$$Pr(S_i = 1|Z_i, \theta_i) = \Phi(Z_i\gamma + \theta_1).$$

Then, given that we do not know θ, we integrate it out and form the likelihood.

$$L = \prod_{i=1}^{N} \int Pr(S_i, Y_i|X_i, Z_i, \theta)\phi(\theta)d(\theta).$$

From this model, we can identify $(\gamma, \beta_0, \beta_1)$ and (α_0, α_1), which can be estimated by maximum likelihood after integrating out θ[5]. Notice that this analysis rests on the normality assumption of θ. For a more flexible functional dependence of the response variable or continuous covariates, see Marra and Radice (2011).

4.1 Sponsorship effect estimation

Given that we can estimate the parameters (γ, α, β), we can estimate the mean and distributional sponsorship effects for the case of a dichotomous outcome variable, as described in Aavik et al. (2005). Let $\Delta = S_1 - S_0$, and $\Delta(x)$ be the expected value of Δ conditioned on $X = x$, then, the sponsorship effect on the population (where the population is all the children within the appropriate age ranges for sponsorship), is defined as $ATE(x) = E[\Delta|X = x]$, and it can be estimated directly by:

$$\begin{aligned} ATE(x) &= Pr(Y_1 = 1|X = x) - Pr(Y_0 = 1|X = x), \\ &= F_{U_1}(x\beta_1) - F_{U_0}(x\beta_0), \\ &= \Phi\Big(\frac{x\beta_1}{\sqrt{1+\alpha_1^2}}\Big) - \Phi\Big(\frac{x\beta_0}{\sqrt{1+\alpha_0^2}}\Big). \end{aligned}$$

Also of interest is to know the impact of sponsorship on the participants in the program, $ATT(x) = E[\Delta|X = x, S = 1]$, which can be estimated by:

$$\begin{aligned} ATT(x,z) &= Pr(Y_1 = 1|X = x, Z = z, S = 1) - Pr(Y_0 = 1|X = x, Z = z, S = 1), \\ &= \frac{1}{Pr(S = 1|X, Z)}(Pr(Y_1 = 1, S = 1|X, Z) - Pr(Y_0 = 1, S = 1|X, Z)), \\ &= \frac{1}{F_{U_S}(z\gamma)}(F_{U_1,U_S}(x\beta_1, z\gamma) - F_{U_0,U_S}(x\beta_0, z\gamma)). \end{aligned}$$

Notice that given the distributional assumptions of the error structure, we can recover the covariates between the unobserved errors to the econometrician, where $Cov(U_S, U_1) = \alpha_1$, $Cov(U_S, U_0) = \alpha_0$, and $Cov(U_1, U_0) = \alpha_0\alpha_1$. Therefore, we can recover the joint distribution of (U_S, U_1, U_0). Furthermore, given the multidimensional normality assumption, we can recover the joint distribution of

[5]We use Gauss-Hermite quadrature with 30 nodes to approximate the integral

(U_S, U_1) and (U_S, U_0), which enables the estimation of the $ATT(x, z)$. Finally, we can also estimate the marginal treatment effect, $MTE(x, u_S)$, which is a building block to estimate both $ATE(x)$ and $ATT(x)$ as shown by Heckman and Vytlacil (1999, 2005, 2007).

The $MTE(x, u_S)$ is defined as $E[\Delta|X = x, U_S = u_S]$, which can be interpreted as the expected effect of sponsorship on the individuals with observed characteristics X, who are indifferent between studying higher education or not, given a value of z such that $z\gamma = u_S$. Notice that for small values of u_S, low levels of $z\gamma$ would be required for the individual to be selected into sponsorship. Similarly, if u_S is large, higher levels of $z\gamma$ would be required for the individual to be selected into sponsorship. In this sense, for low values of u_S, the $MTE(x, u_S)$ can be interpreted as the average effect of sponsorship for the most likely to be selected into the program. Similarly, for high values of u_S, the $MTE(x, u_S)$ would be the average effect of sponsorship for the least likely to be sponsored. The $MTE(x, u_S)$ can be estimated by:

$$\begin{aligned} MTE(x, u_S) &= Pr(Y_1 = 1|X = x, U_S = u_S) - Pr(Y_0 = 1|X = x, U_S = u_S), \\ &= Pr(U_1 \leq x\beta_1|U_S = u_S) - Pr(U_0 \leq x\beta_0|U_S = u_S). \end{aligned}$$

Again, since we have the bivariate distribution of (U_S, U_1) and of (U_S, U_0), and given normality, we also have the distribution of U_1 and of U_0 conditioned on a value of U_S.

5 (Preliminary) Results

Next we report the estimates of the model. We first discuss the estimation of the selection and outcome equation, and then we present the estimation of the mean sponsorship effect parameters. Finally, we further detail the heterogeneity in the impact of sponsorship by presenting the marginal sponsorship effect for different values of u_S.

In the first and third column of Table 3 we observe the coefficients of the selection equation. The results for model 1 correspond to the estimation of the model without the believed return of higher education (ρ_{HE}) in the outcome equation, whereas model 2 does include (ρ_{HE}). Consistent with the selection procedure used by CI, the younger cohorts have higher probability of being selected than older cohorts. Also, being protestant is positively correlated with selection, whereas a higher asset index (greater wealth as measured by household goods) is negatively correlated with selection to the program. In the second and fourth column I report the marginal effect of each covariate defined as $E_Z[\frac{\partial Pr(S=1|Z=z)}{\partial z_k}]$, which is estimated using the analytical derivative averaged over the unconditional distribution of Z. Younger cohorts are around 25 percentage points more likely to be selected into the program, and an increase of one standard deviation in the asset index is related with a decrease of approximately four percentage points of being selected into the program.

The estimates for the outcome equation are presented in Table 4. First to notice is that the coefficients are fairly imprecisely estimated. The standard errors, estimated by bootstrap, are large enough for most of the coefficients to be insignificant. The results suggest that a larger data set may be needed. Given the imprecision of the estimates, these results can be taken as suggestive rather then conclusive. Nevertheless, the signs of the coefficients have the expected direction which

Table 3: Selection equation

	Model 1		Model 2	
	Coeff.	Mg. Effect	Coeff.	Mg. Effect
Factor	1		1	
Age 6	1.702***	0.258***	1.694***	0.257***
	(0.537)	(0.058)	(0.554)	(0.058)
Age 7	1.842***	0.284***	1.846***	0.285***
	(0.51)	(0.052)	(0.406)	(0.048)
Age 8	1.663***	0.27***	1.673***	0.271***
	(0.377)	(0.042)	(0.324)	(0.036)
Protestant	1.924***	0.338***	1.923***	0.338***
	(0.354)	(0.054)	(0.329)	(0.045)
Asset Index	−0.269**	−0.043**	−0.27***	−0.043**
	(0.114)	(0.018)	(0.152)	(0.02)
Treated site	3.161	0.432***	3.149	0.432***
	(2.95)	(0.039)	(3.529)	(0.047)
N	271	271	271	271

we proceed to describe. Under both specifications, the coefficients for the sponsored and the non-sponsored are reported in columns one and three respectively, along with their marginal effects in columns two and four.

Interestingly, when we see the effect of the program *Prospera*, we see that this program is positively correlated with aspirations, both for the sponsored and non-sponsored, but the marginal effect is small. Males have a lower revealed intended choice to acquire higher education both for the sponsored and non-sponsored. Parental education seems to be the most important variable for both sponsored and non-sponsored, with a marginal effect around 5 percentage points for the sponsored, and 2.5 percentage points for the non-sponsored. The asset index is significant and positive only for the non-sponsored, with a marginal effect of 6 percentage points. Finally, we cannot reject the null of no selection on unobservables for the non-sponsored.

When comparing both specifications, the coefficients of the covariates do not change much, both for the sponsored and the non-sponsored, with the exception of the coefficient of the factor (α) for the sponsored, which decreases by 40 percent. We might be concerned that the children may take into account unobservable factors, such as self sense of ability, when answering their believes of their own future income. If that is the case, the elicited returns to schooling would be correlated with the factor and may bias the results. Nevertheless, the inclusion of ρ_{HE} does not have a big effect on the estimation of the sponsorship parameters.

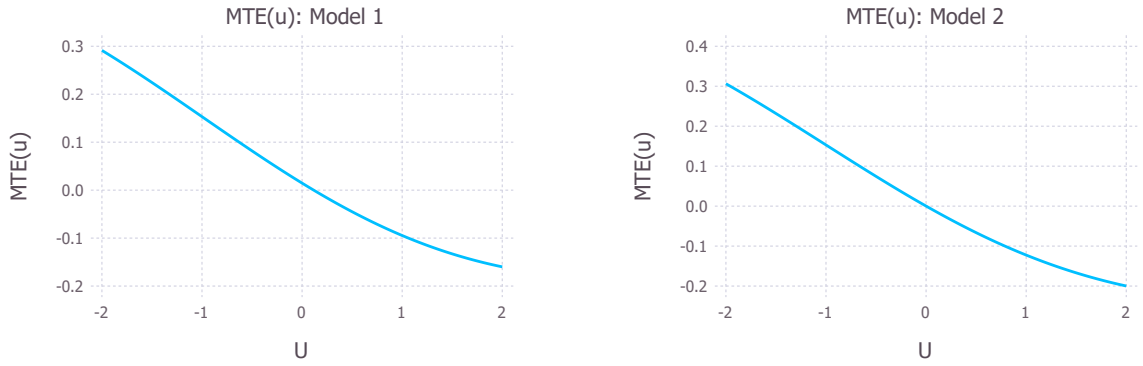
Although imprecisely estimated, due to the large standard deviations, we have that the sponsorship effect on the population is effectively zero, but on the sponsored individuals is about 20 percentage points. Nevertheless, both of these effects are not significant at the standard levels. The difference between the average sponsorship effect, and the sponsorship on the participants can

Table 4: Output equation

	Model1				Model 2			
	Sponsored	Marginal effect	Non-Sponsored	Marginal effect	Sponsored	Marginal effect	Non-Sponsored	Marginal effect
α	-0.137		-0.971^{*}		-0.077		-0.977	
	(1.173)		(0.505)		(0.934)		(0.903)	
Prospera	0.316	0.108	0.005	0.001	0.326	0.112	0.021	0.005
	(0.664)	(0.112)	(0.421)	(0.095)	(0.632)	(0.154)	(0.64)	(0.127)
Parent Educ	0.127	0.048^{***}	0.09^{**}	0.025^{*}	0.138^{*}	0.05^{***}	0.087	0.024
	(0.186)	(0.018)	(0.046)	(0.014)	(0.073)	(0.017)	(0.083)	(0.015)
Male	-0.226	-0.074	-0.14	-0.034	-0.236	-0.078	-0.144	-0.035
	(0.505)	(0.077)	(0.353)	(0.082)	(0.403)	(0.097)	(0.451)	(0.08)
Asset Index	-0.035	-0.011	0.253^{*}	0.061^{***}	-0.033	-0.011	0.25	0.06^{**}
	(0.173)	(0.034)	(0.139)	(0.025)	(0.133)	(0.033)	(0.229)	(0.03)
Distance	-0.001	0	0.003	0.001	-0.001	0	0.003	0.001
	(0.004)	(0.001)	(0.004)	(0.001)	(0.004)	(0.001)	(0.007)	(0.001)
$\rho_{HigherEducation}$					-0.343	-0.113	0.065	0.016
					(0.474)	(0.072)	(0.479)	(0.069)
$E(ATE(x))$	0.006				-0.008			
$V(E(ATE(x)))$	(0.083)				(0.101)			
$E(ATT(x))$	0.199				0.204			
$V(E(ATT(x)))$	(0.183)				(0.180)			
N	271	271	271	271	271	271	271	271

be better explained by the marginal treatment effect (MTE). In Figure 2 we plot the MTE for different values of U_S. The effect of sponsorship is positively related with selection to the program, as the individuals with lower u_S have a higher sponsorship effect. As shown by Heckman and Vytlacil (2007), we can use the MTE as a building block for the ATE and ATT, where the ATE is an integrated version of the MTE using the appropriate weights. Likewise, we can also recover the ATT with higher weights for low levels of u_S, which is consistent with the results obtained. In Figure 3 in the appendix, we graph the MTE using bootstrap, where we graph the mean $MTE(x)$ of the bootstraps along with a 95 percent interval. The mean effect is lower in the bootstrap and has a large variability.

Figure 2: Marginal Treatment Effect



Although the treatment parameters are not significant, the sponsorship effect on the participants is consistent with the results observed by Wydick et al., (2013), who find an impact of CI on completed years of schooling in the range of 1.03 to 1.46 years. If the intended choices would match actual behavior, an increase in 20 percentage points on aspirations would translate into an increase of approximately 8 months of schooling. An interpretation of this result is that even smaller increments in the educational aspirations of children may be valuable, given that aspirations tend to build on themselves.

5.1 Discussion

As noted before, the children may take into account unobservable factors, such as self sense of ability, when answering the questions about their believed future income. In that case the elicited returns to schooling would be correlated with the factor and the results would be biased.

Notice though, that if the believed returns to schooling is correlated with the factor, we can use this variable as a measurement within a measurement system as in Cunha and Heckman (2008). The believed returns is observed for all individuals and does not suffer of a selection problems. Its inclusion in a measurement equation allows us to have three additional covariates, which identify the model through the covariate of the residuals.

A difficulty of estimating this model through the measurement system, is that we would need the estimation of the residuals (U_S, U_1, U_0). The estimation of U_S can be done by using the quantiles of the probit model for the selection equation, and the estimation of U_1 and U_0 is straightforward when we have a continuous outcome. Nevertheless, for the case of binary outcomes we have to account for the nonlinearity of the model. This exercise is the next step in trying to accommodate both the selection problem of the sponsorship procedure and the possible endogeneity in the believed returns to higher education.

6 Conclusions

In this paper we describe the procedure to estimate the impact of the international child sponsorship program *Compassion International* (CI), on the intended choice of schooling of children aged 12 to 15. We first describe the institutional framework of CI, the fieldwork, the measurement of subjective expectations, and the binary Roy type model following Aakvik et al. (2005).

Given the estimation of the parameters of the model, we can further estimate the average effect of the sponsorship program on both the entire population of eligible youth, and on the sponsored individuals. The standard deviation of the outcome coefficients are large, so we see this results as suggestive rather than conclusive. we observe an average effect close to zero and an average effect on the sponsored of around 20 percentage points. Nevertheless, this parameters are not statistically significant at the usual levels. Given our small data set, it seems that we have a problem of statistical power.

The most important variable in our analysis are parental education for both sponsored and non-sponsored, and income measured by an asset index for the non-sponsored. This results suggest that initial conditions seem to play an important role in determining educational aspirations.

7 Bibliography

References

[1] Aakvik, A., J. Heckman, and E. Vytlacil. (2005). Treatment Effects for Discrete Outcomes when Responses to Treatmen Among Observationally Identical Persons an Application to Norwegian Vocational tion Programs. Journal of Econometrics, 125,15-51.

[2] Arcidiacono, Peter (2004). Ability Sorting and the Returns to College Major. Journal of Econometrics 121:343–375.

[3] Arcidiacono, P., Hotz, V.J. and A. Maurel. (2017). Ex Ante Returns and Occupational Choice. Working paper.

[4] Attanasio, O. P. (2009). Expectations and perceptions in developing countries: Their measurement and their use. The American Economic Review 99(2), 87–92.

[5] Attanasio, O. P. and K. M. Kaufmann (2009). Educational choices, subjective expectations and credit constraints. NBER Working Paper 15087 July.

[6] Attanasio, O.P., Meghuir C., and A. Santiago (2012) Education Choices in Mexico: Using a Structural Model and a Randomized Experiment to Evaluate PROGRESA. The Review of Economics and Statistics. 79(1):37-66.

[7] Bajari, P. (1997). Econometrics of the first-price auctions with asymmetric bidders, mimeo, Harvard University.

[8] Beaman, L., E. Duflo, R. Pande and P. Topalova. 2012. Female Leadership Raises Aspirations and Educational Attainment for Girls: A Policy Experiment in India. Science Magazine, 335(6068): 582-586.

[9] Bernheim, D. (1988): "Social Security Benefits: An Empirical Study of Expectations and Realizations," in Issues in Contemporary Retirement, ed. by R. Campbell and E. Lazear. Stanford: Hoover Institution.

[10] Cunha, F., and J. Heckman. (2008). A New Framework for the Analysis of Inequality. Macroeconomic Dynamics, Vol 12(S2): 315-354.

[11] Cunha, F., and J. Heckman. (2009). The Economics and Psychology of Inequality and Human Development. Journal of the European Economic Association, Vol. 7(2-3):1-48.

[12] Carneiro, P., Heckman, J. J., and Vytlacil, E. J. (2011). Estimating Marginal Returns to Education. American Economic Review, 101(October), 2754–2781.

[13] Dalton, P, Sayantan G, and A Mani. (2016). Poverty and Aspirations Failure: A Theoretical Framework. Economic Journal, 126(590):165-188.

[14] Delavande, A., X. Giné, and D. McKenzie (2011). Measuring subjective expectations in developing countries: A critical review and new evidence. Journal of Development Economics 94(2), 151–163.

[15] Delavande, A., and B. Zafar. (2017) Intended College Attendance: Evidence from an Experiment on College Returns and Costs. Forthcoming, Journal of Public Economics.

[16] Dominitz, J., and Manski (1996): "Eliciting Student Expectations of the Returns to Schooling," Journal of Human Resources, 31, 1–26. (1997a): "Using Expectations Data to Study Subjective Income Expectations," Journal of the American Statistical Association, 92, 855–867.

[17] Dominitz, J., and Manski (2003): How Should We Measure Consumer Confidence (Sentiment)? Evidence from the Michigan Survey of Consumers, Working Paper 9926, National Bureau of Economic Research.

[18] Dominitz, J., and Manski (2004): How Should We Measure Consumer Confidence? Journal of Economic Perspectives, 18(2):51-66.

[19] Filmer, D., and L. Pritchett (2001). Estimating Wealth Effects without Expenditure Data-or Tears: An Application to Educational Enrollments in States of India. Demography, 38(1):115-132

[20] Glewwe, Paul, Phillip. H. Ross, and Bruce Wydick. (2017) Developing Hope Among Impoverished Children: Using Child Self-Portraits to Measure Poverty Program Impacts, forthcoming, Journal of Human Resources.

[21] Gustman, A., and T. Steinmeier (1999): What People Don't Know about Their Pensions and Social Security: An Analysis Using Linked Data from the Health and Retirement Study, National Bureau of Economic Research Working Paper 7368.

[22] Gustman, A., and T. Steinmeier (2001): Imperfect Knowledge,Retirement, and Savings, National Bureau of Economic Research Working Paper 8406.

[23] Heckman, J. J. and T. Kautz (2012). Hard evidence on soft skills. Labour Economics. 19(4): 451–464.\{}bibitem{dz1} Delavande, A., and B. Zafar. (2017) Intended College Attendance: Evidence from an Experiment on College Returns and Costs. Forthcoming, Journal of Public Economics.

[24] Heckman, J. J., R. Pinto, and P. A. Savelyev (2013). Understanding the mechanisms through which an influential early childhood program boosted adult outcomes. American Economic Review, 103(6): 1–35.

[25] Heckman, J., and E. J. Vytlacil. (1999). Local Instrumental Variables and Latent Variable Models for Identifying and Bounding Treatment Effects. Proceedings of the National Academy of Sciences, 96(8): 4730–34.

[26] Heckman, J.J. and E. Vytlacil (2005). Structural equations, treatment effects, and econometric policy evaluation. Econometrica, 73 (3), 669–738.

[27] Heckman, James J., and Edward J. Vytlacil. (2007). "Econometric Evaluation of Social Programs, Part II: Using the Marginal Treatment Effect to Organize Alternative Econometric Estimators to Evaluate Social Programs, and to Forecast their Effects in New Environments." In Handbook of Econometrics, Vol. 6B. Handbooks in Economics 2, ed. James J. Heckman and Edward E. Leamer, 4875–5144. New York: North-Holland.

[28] Jacob, B., and T. Wilder. (2011). "Educational Expectation and Attainment." In Whither Opportunity? Rising Inequality and the Uncertain Life Chances of Low-Income Children, edited by Greg J. Duncan and Richard J. Murnane. New York, NY: Russell Sage Press.

[29] Jensen, R. (2010). The (perceived) returns to education and the demand for schooling. Quarterly Journal of Economics, 125(2), 515–548.

[30] Kaufman, K.M. (2014). Understanding the Income Gradient in College Attendance in Mexico: The Role of Heterogeneity in Expected Returns. Quantitative Economics, 5(3):583-630

[31] David McKenzie (2005) Measuring inequality with asset indicators. Journal of Population Economics, 18(2):229–260

[32] Manski, C.F. (1990): The Use of Intentions Data to Predict Behavior: A Best Case Analysis, Journal of the American Statistical Association, 85, 934–940.

[33] Manski, C.F. (1993): Adolescent Econometricians: How Do Youth Infer the Returns to Schooling? in Studies of Supply and Demand in Higher Education, ed. by C. Clotfelter and M. Rothschild. Chicago: University of Chicago Press, 43–57.

[34] Manski, C.F. (2004). Measuring expectations. Econometrica 72 (5), 1329–76.

[35] Marra, G. and Radice, R. (2011). Estimation of a semiparametric recursive bivariate probit model in the presence of endogeneity. Canadian Journal of Statistics, 39 259–279.

[36] Nguyen, T. (2008). Information, role models and perceived returns to education: Experimental evidence from Madagascar. mimeo

[37] Rosenberg, Morris. 1965. "Society and the Adolescent Self-Image," Princeton, NJ:Princeton University Press.

[38] Stinebrickner, R. and T. R. Stinebrickner. (2012). Learning about Academic Ability and the College Dropout Decision. Journal of Labor Economics. 30(4):707-748.

[39] Stinebrickner, R. and T. R. Stinebrickner. (2014). Academic Performance and College Dropout: Using Longitudinal Expectations Data to Estimate a Learning Model. Journal of Labor Economics. 32(3):601-644.

[40] Todd, P. and Kenneth Wolpin. (2006) Assessing the Impact of a School Subsidy Program in Mexico: Using a Social Experiment to Validate a Dynamic Behavioral Model of Child Schooling and Fertility. American Economic Review, Vol. 96(5):1384-1417.

[41] Wiswall, M., and B. Zafar. (2015) Determinants of College Major Choice: Identification Using an Information Experiment. Review of Economic Studies. 82(2):791-824.

[42] Wydick, B., P. Glewwe and Rutledge, L. (2013). Does International Child Sponsorship Work? A Six-Country Study of Impacts on Adult Life Outcomes. Journal of Political Economy, 121(2): 393-426.

[43] Zafar, B. (2013). College Major Choice and the Gender Gap. Journal of Human Resources. 48(3):545-595.

8 Appendix

Table 5: Summary statistics: sponsored vs non-sponsored

	Site with CI			Site without CI	
	Sponsored	Sibling of sponsored	Non-sponsored in non-sponsored HH	Non-sponsored in non-sponsored HH	Mean all
Higher educ	0.679	0.667	0.838	0.705	0.731
	(0.469)	(0.483)	(0.371)	(0.460)	(0.444)
University	0.550	0.571	0.725	0.607	0.616
	(0.500)	(0.507)	(0.449)	(0.493)	(0.487)
Age	13.046	14.333	13.663	13.328	13.391
	(1.013)	(0.796)	(1.102)	(1.076)	(1.100)
Male	0.413	0.524	0.463	0.492	0.454
	(0.495)	(0.512)	(0.502)	(0.504)	(0.499)
Birth order	2.430	1.857	2.788	2.431	2.492
	(1.632)	(0.854)	(2.085)	(1.874)	(1.798)
Num. siblings	4.055	5.000	3.938	3.719	4.022
	(2.085)	(1.703)	(2.071)	(1.730)	(1.996)
Protestant	0.798	0.524	0.388	0.246	0.531
	(0.403)	(0.512)	(0.490)	(0.434)	(0.500)
Prospera	0.862	0.905	0.850	0.836	0.856
	(0.346)	(0.301)	(0.359)	(0.373)	(0.352)
Education father	7.085	7.158	6.724	6.593	6.873
	(3.193)	(3.270)	(3.289)	(3.420)	(3.267)
Education mother	6.101	5.190	6.987	6.852	6.461
	(3.391)	(3.487)	(2.853)	(3.395)	(3.277)
Asset index	-0.217	-0.125	0.389	0.301	0.085
	(1.294)	(1.486)	(1.190)	(1.503)	(1.352)
N	109	21	80	61	271

Table 6: Comparisons by groups

	(Sponsored) vs (Non-sponsored in sponsored sites)	(Sponsored) vs (Non-sponsored in non-sponsored sites)	(Non-sponsored in non-sponsored site) vs (Non-sponsored in sponsored site)
	t-test	t-test	t-test
Higher educ	-0.159**	-0.026	-0.144**
	(0.062)	(0.073)	(0.071)
University	-0.175**	-0.056	-0.128
	(0.070)	(0.078)	(0.079)
Age	-0.617***	-0.282*	-0.340*
	(0.150)	(0.164)	(0.174)
Male	-0.050	-0.079	0.021
	(0.073)	(0.080)	(0.085)
Birth order	-0.358	-0.001	-0.347
	(0.319)	(0.299)	(0.368)
Num. siblings	0.118	0.336	-0.162
	(0.353)	(0.328)	(0.357)
Protestant	0.411***	0.552***	-0.129
	(0.072)	(0.070)	(0.083)
Prospera	0.012	0.026	-0.011
	(0.055)	(0.065)	(0.070)
Education father	0.361	0.492	-0.190
	(0.521)	(0.615)	(0.656)
Education mother	-0.887*	-0.752	-0.246
	(0.490)	(0.584)	(0.593)
Asset index	-0.606***	-0.518**	-0.135
	(0.200)	(0.261)	(0.261)
N	189	170	142

Figure 3: Marginal Sponsorship Effect: Higher Education

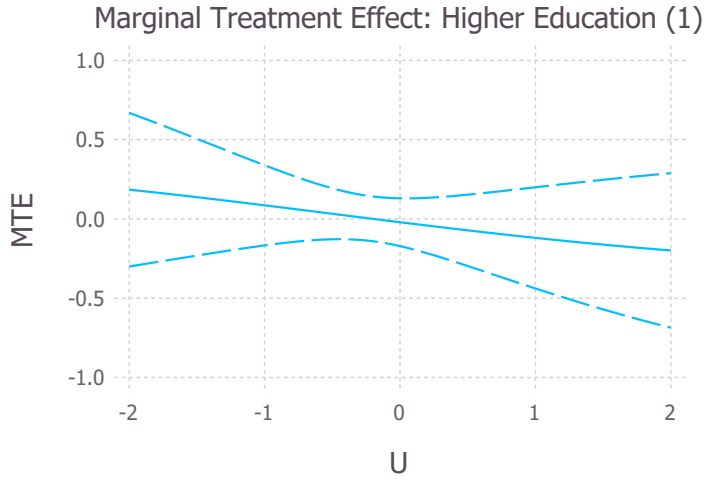


Table 7: Comparisons by site

	Mean all (std. dev.)	Sponsored site (std. dev.)	Non-sponsored site (std. dev.)	t-test (std. dev.)
Higher educ	0.731	0.742	0.694	0.048
	(0.444)	(0.439)	(0.465)	(0.066)
University	0.616	0.622	0.597	0.025
	(0.487)	(0.486)	(0.495)	(0.070)
Age	13.391	13.411	13.323	0.089
	(1.100)	(1.111)	(1.068)	(0.149)
Male	0.454	0.445	0.484	-0.039
	(0.499)	(0.498)	(0.504)	(0.072)
Birth order	2.492	2.507	2.441	0.067
	(1.798)	(1.784)	(1.859)	(0.282)
Num. siblings	4.022	4.091	3.776	0.315
	(1.996)	(2.054)	(1.768)	(0.287)
Protestant	0.531	0.612	0.258	0.354***
	(0.500)	(0.488)	(0.441)	(0.068)
Prospera	0.856	0.861	0.839	0.023
	(0.352)	(0.347)	(0.371)	(0.060)
Education father	6.873	6.975	6.533	0.442
	(3.267)	(3.221)	(3.422)	(0.569)
Education mother	6.461	6.378	6.742	-0.364
	(3.277)	(3.219)	(3.478)	(0.536)
Asset index	0.085	0.035	0.254	-0.219
	(1.352)	(1.293)	(1.535)	(0.241)
N	271	209	62	271

Table 8: Expected income vs census data data at age 24 (mean of data)

	High-school			Techical school			University		
	Male	Female	Total	Male	Female	Total	Male	Female	Total
Census									
Oaxaca-less50k	4829.2	3898.4	4484.8	5693.3	4646.7	5102.9	6446.6	6162.8	6291.9
	(4293.9)	(2577.1)	(3777.6)	(3296.6)	(2939.9)	(3123.7)	(3661.4)	(3058.3)	(3348.0)
Oaxaca-50k	5434.2	4313.6	4971.7	7933.1	4485.5	5470.5	5680.2	6183.7	5948
	(4359.6)	(1691.5)	(3549.4)	(5394)	(2008.4)	(3220)	(3052.2)	(3231.6)	(3151.8)
Chiapas-less50k	3681	3189.4	3561.5	4940.9	3971.4	4545.9	5770.6	5506.5	5650.9
	(2749.4)	(2300)	(2655)	(4208.7)	(2646.5)	(3656.1)	(3504.8)	(3042)	(3304.2)
Chiapas-50k	4685.3	3867.2	4374	7847.8	5307.6	6718.8	8356.8	6567.6	7496.2
	(2444.1)	(2397.4)	(2456.6)	(4876.8)	(3806.1)	(4372.2)	(27981.3)	(4658.4)	(20422.7)
Survey									
Oaxaca	6270	7200.5	6743.7	9676.9	8908.4	9285.7	13562.8	11366.3	12444.6
	(4772.6)	(14176.3)	(10567.5)	(7993.7)	(15996.3)	(12603.9)	(11293.2)	(19883.3)	(16134.1)
Chiapas	5531.5	4967.1	5217.9	8421.1	7027.1	7646.7	13365.6	11406.7	12277.3
	(5897.1)	(4347.2)	(5089.7)	(8761.4)	(5886.5)	(7319.8)	(14966.2)	(11228.3)	(13027)
AK(2014)-averages	3525.5	2509.3					6110.6	4806.8	

Table 9: Expected income vs census data data at age 24 (log of data)

	High-school			Techical school			University		
	Male	Female	Total	Male	Female	Total	Male	Female	Total
Census									
Oaxaca-less50k	8.29	8.07	8.21	8.49	8.24	8.35	8.62	8.59	8.6
	(0.61)	(0.68)	(0.65)	(0.63)	(0.67)	(0.66)	(0.59)	(0.57)	(0.58)
Oaxaca-50k	8.43	8.28	8.37	8.85	8.29	8.45	8.52	8.61	8.57
	(0.54)	(0.46)	(0.51)	(0.74)	(0.62)	(0.65)	(0.49)	(0.49)	(0.49)
Chiapas-less50k	7.97	7.87	7.95	8.21	8.1	8.16	8.48	8.44	8.46
	(0.81)	(0.64)	(0.77)	(0.85)	(0.6)	(0.76)	(0.65)	(0.64)	(0.65)
Chiapas-50k	8.35	8.15	8.27	8.88	7.85	8.63	8.57	8.57	8.57
	(0.45)	(0.51)	(0.48)	(0.75)	(NA)	(0.8)	(0.66)	(0.55)	(0.61)
Survey									
Oaxaca	8.33	8.12	8.22	8.8	8.43	8.61	9.13	8.71	8.91
	(1.03)	(1.1)	(1.06)	(0.95)	(1.03)	(1)	(0.96)	(1)	(0.99)
Chiapas	8.23	8.19	8.21	8.62	8.55	8.58	9.03	8.99	9.01
	(0.9)	(0.79)	(0.84)	(0.93)	(0.78)	(0.85)	(0.98)	(0.81)	(0.88)
AK(2014)	7.58	8.13					7.24	7.89	

Table 10: Expected income vs census data at age 24 (median of data)

	High-school			Techical school			University		
	Male	Female	Total	Male	Female	Total	Male	Female	Total
Census									
Oaxaca-less50k	4271.7	3295.6	4119	5186.9	4271.7	4577.1	6407.6	6407.6	6407.6
	(4293.9)	(2577.1)	(3777.6)	(3296.6)	(2939.9)	(3123.7)	(3661.4)	(3058.3)	(3348.0)
Oaxaca-50k	4577.1	3867.5	4271.7	7933.1	4577.1	4577.1	4698.9	5339.6	5339.6
	(4359.6)	(1691.5)	(3549.4)	(5394)	(2008.4)	(3220)	(3052.2)	(3231.6)	(3151.8)
Chiapas-less50k	3203.8	2745.6	3203.8	3753.2	2867.9	3661.9	5339.6	4577.1	4752.3
	(2749.4)	(2300)	(2655)	(4208.7)	(2646.5)	(3656.1)	(3504.8)	(3042)	(3304.2)
Chiapas-50k	4271.7	3547.1	3844.5	6407.6	2563	5034.7	5339.6	5339.6	5339.6
	(2444.1)	(2397.4)	(2456.6)	(4876.8)	(3806.1)	(4372.2)	(27981.3)	(4658.4)	(20422.7)
Survey									
Age-12	4735.9	3205.4	3832.9	6467.9	4740.2	5546.5	9468.2	7442.2	7577.5
	(4392.3)	(4283.1)	(4348.7)	(5646.7)	(5830.4)	(5772.4)	(8438.8)	(13656.1)	(11589.2)
Age-13	3449	3144.4	3409.9	5177.3	5180	5180	6820.8	7417.1	7417.1
	(9069.9)	(4906.1)	(6707.5)	(12058.4)	(6473)	(8940.7)	(15889.2)	(10843.9)	(12901.5)
Age-14	3656.4	4041.3	3716.6	5277.2	6312.5	5352.4	7059	8162.9	7417.1
	(4155.7)	(3329)	(3773.4)	(8156.4)	(5366.7)	(6958.8)	(13691.5)	(8169.8)	(11398.7)
Age-15	4484.6	3226.9	3792.8	6701.2	4480.3	5538.1	11772.2	7417.1	9461.3
	(4012.4)	(13951.1)	(10204.7)	(7904.2)	(15574.5)	(12242.3)	(17464.6)	(19367.6)	(18397.2)
AK(2014)-averages	3525.5	2509.3					6110.6	4806.8	

Table 11: Expected income vs census data ages 24 to 65 (median of data)

	High-school			Techical school			University		
	Male	Female	Total	Male	Female	Total	Male	Female	Total
Census									
Oaxaca-less50k	4577.1	4271.7	4577.1	6407.6	5492.4	6194	8238	7475.5	7902.7
	(4860.0)	(3601.9)	(4495.5)	(5822.1)	(4114.8)	(4993.3)	(9911.6)	(4636.2)	(7900.6)
Oaxaca-50k	5492.4	4577.1	5339.6	6865.7	6407.6	6407.6	8543.4	7475.5	8238
	(4918.7)	(3971.8)	(4654.6)	(8123.1)	(7940.6)	(8041.9)	(7581.5)	(5521.9)	(6751.4)
Chiapas-less50k	3432.3	3203.8	3203.8	4698.9	4698.9	4698.9	7475.5	7475.5	7475.5
	(4274.8)	(3177.2)	(4068.2)	(44577.2)	(3963.5)	(32165.1)	(6861.4)	(5150.1)	(6283.4)
Chiapas-50k	4577.1	3844.5	4271.7	6407.6	5339.6	5339.6	8543.4	7475.5	8543.4
	(10178.3)	(3635.6)	(8623.2)	(5718.7)	(3967.3)	(4753.4)	(17923.2)	(6475.8)	(14226.7)
Survey									
Age-12	4735.9	3205.4	3832.9	6467.9	4740.2	5546.5	9468.2	7442.2	7577.5
	(4392.3)	(4283.1)	(4348.7)	(5646.7)	(5830.4)	(5772.4)	(8438.8)	(13656.1)	(11589.2)
Age-13	3449	3144.4	3409.9	5177.3	5180	5180	6820.8	7417.1	7417.1
	(9069.9)	(4906.1)	(6707.5)	(12058.4)	(6473)	(8940.7)	(15889.2)	(10843.9)	(12901.5)
Age-14	3656.4	4041.3	3716.6	5277.2	6312.5	5352.4	7059	8162.9	7417.1
	(4155.7)	(3329)	(3773.4)	(8156.4)	(5366.7)	(6958.8)	(13691.5)	(8169.8)	(11398.7)
Age-15	4484.6	3226.9	3792.8	6701.2	4480.3	5538.1	11772.2	7417.1	9461.3
	(4012.4)	(13951.1)	(10204.7)	(7904.2)	(15574.5)	(12242.3)	(17464.6)	(19367.6)	(18397.2)
AK(2014)-averages	3525.5	2509.3					6110.6	4806.8	