The Effects of Subsidizing Secondary Schooling: Evidence From a Conditional Cash Transfer Program in Costa Rica

Jaime A. Meza-Cordero

University of Southern California

Abstract: In 2006 the Costa Rican government introduced a nationwide conditional cash transfer program under the name "Avancemos". This program was based on giving students' parents a monthly subsidy, conditional on mandatory school attendance for each child. The program was introduced prioritizing those children in greater danger of dropping out of school; it was not a randomized experiment. In this paper I aim to estimate the effects of the program on years of school completed by using a difference-in-differences propensity score matching technique. I find that after 4 years in place, the subsidy increases the years of school completed of eligible students in about half a year.[*]

August 2011

[*] I thank the Costa Rican National Institute of Census and Statistics (INEC) for access to their databases. I would also like to thank Leonardo Garnier and José Antonio Li for their valuable information about Avancemos.

I Introduction

In 2005 the attendance rate for high school in Costa Rica had fallen to one of its all-time lows at 78.4%, with only 20% of the students finishing the five years of high school on time. It is well know that not finishing secondary school leads to poverty, and at the same time poverty leads to underage employment in the informal sector, which in turn leads to even lower schooling. As an attempt to alleviate this problem, in 2006 the Costa Rican Government created the "Avancemos" Program. The program consisted on a monetary subsidy conditional on school attendance and a yearly health check for eligible adolescents. These types of programs are popularly known as conditional cash transfers (from now on CCT). In the last decades CCT programs have become very widespread, especially in Latin America.

Human capital plays a very important role in economic development, the higher the education levels, the higher the productivity and the higher the wages received, among other incomes. Having a well-educated labor force also facilitates the adoption and efficient use of the most modern technologies. There is a generalized belief that education years have decreasing returns, this leads to think that it is very important to retain students in school during their early ages. Also, it is well known that low schooling and poverty are positively correlated, so in order to fight back the poverty cycle it becomes very important to stimulate social mobility by providing schooling opportunities to the poorest children.

CCT programs should provide short-term financial relief to participating households and lead to long-term wealth redistribution through human capital. Since the poorest population has the lower access to risk reduction mechanisms, they are the most vulnerable during economic downturns. When an economic crisis arises the poorest families usually react by forcing their children to help raising income for the family, leading to lower schooling and intergenerational poverty. To address this issue I intend to analyze the causes and evolution of dropout rates.

The expected effects of this program are: a short-term increase in attendance rates, a short-term positive effect on years of school completed, a short-term decrease in under-age hours worked, a long-term decrease in poverty, and an ambiguous effect on grades and learning due to the positive effect associated with higher school attendance offset by the negative effect of a larger classroom size. Also, I would expect an increase in consumption goods and services, better nutrition and better health due to the increase in disposable income within the household. According to the Minister of Education, Leonardo Garnier, high school attendance increased from 78.4% in 2005 to 82.9% in 2009, confirming the expected short-term positive effect on attendance rates. In this paper my main objective is to determine the magnitude of the changes in years of school completed after the first 4 years of the program's existence.

The policy change in 2006 suggest the need for a before and after policy evaluation. However, comparing average results before and after the program is not a good design due to the existence of time specific effects. Therefore, I will make use of a mean-difference analysis after the program was implemented in which I will compare the treated group with a similar group that would qualify for the program was not selected. This control group will be created to match observable characteristics with the treated group.

The data for the empirical analysis comes from yearly household surveys, which contain information of numerous socioeconomic characteristics as well as the main variable of interest: receiving the conditional cash transfer or not. Since this program is very recent, in this study I can only use data for up to 4 years after the program started. However, as years pass by I have the intention of updating the data for analyzing longer-term results.

The structure of this paper is the following: A brief description of the Avancemos Program, a review of some of the existing literature on CCT programs, a description of the data, the theoretical framework and empirical strategy used. And finally, a presentation of the obtained results and my concluding remarks.

II The Avancemos Program

In mid 2006 the CCT project started with a pilot plan intended to develop how to run the program effectively. This plan took place only in underprivileged regions located in urban areas and covered 8000 high school students. Starting in 2007 the program expanded nationally as regional quotas were determined based on each County's Index of Social Development, giving more resources to the regions with a lower index. The program's coverage kept growing every year: 8,000 students in 2006, 94,621 in 2007, 136,000 in 2008, 165,749 in 2009 and 185,000 for 2010 and 2011; which is about 51.4% of the currently enrolled students.

The Avancemos program has the main objective to offset the potential income a student could get from dropping out of school and working. The monthly transfer was designed to rise with school year due to the higher opportunity cost associated with age. For 7th grade the payment was of around $30, $40 for 8th, $50 for 9th, $70 for 10th and $80 for 11th. The jump in 10th grade is due to the fact that the dropout rate after 9th grade is especially high. To obtain an idea of the magnitude of these transfers, I should point out that the minimum wage at that time was of about $400. Although the program doesn't forbid having a job, it at the very least limits the time available for working, and it is expected that a student not working would have better school attendance and higher grades. For this program only high school students were eligible and all households were uniformly notified of the program's implementation through national media. There is also no restriction on the number of eligible children per family as long as they don't have repeated a grade more than once.

Regarding eligibility and selection, a study made by Molina & Fallas (2009) defines as requisites for the students: to be qualified as in social exclusion, registered in high school, undergoing a yearly heath check, and submitting a contract signed by a parent or legal guardian. They also explain how based on income, each family is fitted in one of three poverty level groups: not poor, poor, or extremely poor. Given the poverty level group, school recommendations, and community recommendations a score level is computed for each student, the ones with the highest score are selected as beneficiaries for the CCT.

To confirm these norms I interviewed Jose Antonio Li, who is the former president of the institution in charge of the program, he explained me how the selection criteria is based on three steps. First, the family has to be identified by the IMAS (Mixed Institute of Social Aid), and must have filled a technical survey named FIS that contains extensive information regarding the socioeconomic status of the family throughout 56 variables which are corroborated in situ. The poorest families (about a third of the country) are covered through these surveys and they are categorized into 4 groups depending on their socioeconomic status. A second step consists on a parent applying for the program with a signed contract. In the last step, the IMAS establishes a regional cutoff point under which all students are selected for the subsidy. After the students are selected, the money is given out in a monthly basis to the mothers through their bank accounts.

III Literature Review

School desertion is a very common problem in underdeveloped countries, thus CCT programs come to play a very important role in education. Several authors have acknowledged the importance of these programs and have tried to measure their effects. Parker et al. (2008) point out: "The main motivation of conditional cash transfer programs is the linking of benefits to human capital investment, particularly of children. The aim is to alleviate current poverty through monetary transfers as well as future poverty, by increasing human capital of children". There are numerous studies that have tried to determine the impacts of CCT programs, however, most of the programs evaluated were developed as randomized experiments.

The most studied CCT program to this day is the Progresa/Oportunidades program in Mexico, which started as a randomized experiment in 1998. The program included elementary and high school students who lived in rural areas. Schultz (2004) calculated the increase in the probability of finishing another year of schooling, conditioned on having finished the previous year. His results are that the program increases schooling in 0.81 years using a single difference approach and 0.66 years using a more appropriate difference-in-difference approach. Behrman, Sengupta & Todd (2001) found that the program increased grades completed in 0.6 years and school attendance in 19%. Skoufias, Davis and Behrman (1999) found an increase of 10% in school attendance.

In a set of newer studies, Behrman et al. (2006) and Parker et al. (2008) explain how Progresa was renamed Oportunidades as it extended to urban Mexico. The program now began to work based on eligibility requirements and self-selection because individuals were required to apply for the program in a module. In this module socioeconomic status was identified and later verified by an in situ visit. Behrman et al. (2006) provide estimates of the program's effect on enrollment and grade completion based on a difference-in-difference matching approach. Parker et al. (2008) also used matching to create comparison groups. Both studies find positive significant effects on school enrollment and grade completion. This second stage of the Progresa/Oportunidades program is very similar to the Avancemos program, that's why I'm also going to use matching methods in the empirical part of this paper.

Another big CCT program in Latin America is Brazil's Bolsa Escola, which started in 1995 and expanded in 2001. The subsidy was given to families with children between ages 6-15 and who had a per capita income lower than R$90. Costa Resende & De Oliveira (2008) found an increase in food consumption (quality, diversification and quantity).

Colombia also developed a CCT program named Familias en Acción. The program started in 1999 and it had two components: From ages 0-7 there was a monetary transfer to the mothers intended to improve the nutrition of the child, this conditional on medical check ups. And from ages 7-17 there was a monetary transfer to parents conditional on keeping the student(s) in school. Barrero et al. find an increase of 12.7% in returns from education.

The first CCT program in Costa Rica was the Superémonos Program, which was very small scale. This program took place in the early 2000's and gave food subsidies to mothers conditional on school assistance. As in Avancemos, random assignment was not done, for this reason Duryea and Morrison (2004) use non-experimental techniques for it's evaluation. The authors try to find the impact of the program on school attendance, school performance and child labor by using 3 different analyses: mean differences, regression and propensity score matching. They find significant increases in attendance but no effects on performance and child labor.

As there are several defenders of CCT programs, there are also diverse authors who have been very critical about the outcome of these programs. Ravallion & Wodon (1999) study a non-experimental program in which rice is given conditioned on school attendance in Bangladesh. They argue that this conditional transfer program led to lower leisure for kids due to the fact that the parents forced them to stay working. On another study, Skoufias & Parker (2001) claim that CCT are not necessarily the best way to spend money in the educational system, they claim that supply conditions should be also considered (investing in school infrastructure, better professors, etc.).

IV Data: The Costa Rican Household Survey

The data set used is the Costa Rican Household Survey for Multiple Purposes (Encuesta Nacional de Hogares con Propositos Multiples). This survey consists of yearly cross sections, which follow the same methodology since 1999. The primary unit is the household and the sample was chosen from 826,541 units. The sample is stratified in order to have a good representation of all the regions in the country. In a first stage clusters of households are selected, and in a second stage individual households are selected randomly. The sample consists of 10,890 households.

Although this data set is very complete, several limitations arise. The first limitation comes due to the omission of non-monetary factors such as informal work and work inside the house (raising children, cooking, constructing, etc), which are not considered. Each year has a different sample of households, so panel data approaches for individuals can't be used. Finally, since the program is only 4 years old, only short-term effects such as school completion can be estimated, long-term effects such as income distribution and poverty reduction can't be studied.

As I mentioned earlier, the program started with a pilot plan in 2006, but it was in 2007 when it expanded nationwide. In Graph 1 in the appendix I show how average years of schooling for adolescents between ages 12 and 17 (the eligible age group) were increasing very slowly in the first part of the decade and how after 2007 there has been a substantial increase. Graph 2 shows how after 2007 hours worked by adolescents have decreased. From these plots we can observe an important behavioral change in adolescents after the program started, in what follows I'm going to try to quantify these effects in the selected students.

One of the objectives of this paper is to analyze the evolution of high school students' attendance; in Table 1 I present information from the set of eligible students in the years 2000, 2005 and 2010. It's important to highlight that every year more students are attending high school, so there's a time trend even before the Avancemos Program started. It's also important to notice that average grade completion has also increased in time for every age.

Another objective of this study is to understand why there have been so many dropouts in the past decade; in Table 2 I detail the reasons to quit given by the students who have dropped school. Graph 4 shows the evolution of the main reasons for dropouts in the last 10 years. Not being interested in formal education, can't affording studies, and having to work come as the main reasons for quitting school. Having to work is especially important because the CCT is set to offset this reason. In graph 5 it can be seen how quitting school due to the "Have to work" reason has decreased after 2007. Graph 6 depicts the decrease in adolescents having a primary job after 2007. As well as having to work, affording school is very linked to receiving the subsidy or not. It can be seen in Table 2 how this variable significantly decreased as a cause for dropping out in 2010, very likely due to the CCT program installment. It is interesting to notice that not being interested in formal education has remained consistently throughout the years as the main reason for dropouts.

Another important topic is the composition of the eligible students. From the baseline survey I find that within adolescents ages 12-17, about 18.7% don't receive any education, while 56.6% are registered in high school and 21.9% in elementary school. This last 21.9% is mostly composed of students who had to repeat grades, 65.2% of them are age 12 and 22.7% are age 13. Another check I made was on the selected students' parents. I looked for a behavioral change in the head of the household's working hours due to the subsidy now in place. I find that there's no significant change in hours worked, the results are presented in Table 3.

Finally, I wanted to know how well distributed the subsidy was among the poorest families based on some characteristics of those who were selected in 2007. I find that out of the students who get the program only 16.2% have a family car, only 20.9% have a computer, and almost 100% of these students go to public schools. Thus, it seems that there's a low level of filtration. A complete set of summary statistics is presented in tables 4 and 5.

V Theoretical Framework

The Avancemos program was not a randomized experiment. Since post-program differences between treated and untreated groups could come from pre-program differences (due to selectivity), then comparing means between groups with and without the subsidy would not reveal the true effects of the program. To avoid this problem I need to generate an ideal comparison group in order to estimate program effects by comparing weighted means between treated and untreated groups.

I will use the Propensity Score Matching (PSM) technique (Rosenbaum & Rubin 1983), which generates a control group based on observable characteristics. Nowadays this method is very popular due to its easy application and diverse advantages. According to Ravallion (2008): "In evaluating anti-poverty programs in developing countries, single-difference comparisons using PSM have the advantage that they do not require either randomization or baseline (pre-intervention) data." Since the data used to determine cut-off points is collected by the IMAS and is not publicly available, and the data I have comes from household surveys, approaches such as regression discontinuity can't be correctly performed.

The treatment variable will be the dummy D that is equal to 1 if receiving the CCT and 0 if not. The dependent variable will be years of school completed, it's defined as Y_{1i} if receiving the CCT and Y_{0i} if not. I am mostly interested in the average treatment effect on treated (ATT_i), which is:

$$ATT_i = E(Y_{1i} - Y_{0i} \mid D_i=1) = E(Y_{1i} \mid D_i=1) - E(Y_{0i} \mid D_i=1) \qquad (1)$$

The average treatment effect (ATE_i) of the program would be:

$$ATE_i = E(Y_{1i} - Y_{0i} \mid D_i=1)Pr(D=1) + E(Y_{1i} - Y_{0i} \mid D_i=0)Pr(D=0) \qquad (2)$$

Since an eligible individual i can't be simultaneously in the treatment and the control group at the same time, then $E(Y_{0i} \mid D_i=1)$ (the counterfactual) doesn't exist. Therefore, I have to assume that there's a similar group of non-participants, such that: $E(Y_{0i} \mid D_i=1) = E(Y_{0i} \mid D_i=0)$. There would not be selection bias if randomized placement were used, however, that's not the case for Avancemos. I have to deal with a non-experimental evaluation in which I expect a considerable amount of selection bias due to the existence of unobservables such as: distance to the school and public administration offices, student's motivation and expectations of the parents. This selection bias can be seen from a simple single difference in mean outcomes between treated and untreated equation:

$$Diff(Y) = E(Y_1 | X, D=1) - E(Y_0 | X, D=0)$$

$$Diff(Y) = [E(Y_1 | X, D=1) - E(Y_0 | X, D=1)] + [E(Y_0 | X, D=1) + E(Y_0 | X, D=0)]$$

Diff(Y) = Average Treatment Effect on the Treated + Selection Bias

To account for this problem I make use of the observable characteristics in selection "X":

$$ATT_i = E(Y_{1i} - Y_{0i} \mid D_i=1, X) = E(Y_{1i} \mid D_i=1, X) - E(Y_{0i} \mid D_i=1, X) \qquad (3)$$

At this point I make use of the identification hypothesis: In selection, people with the same characteristics have the same chance of being selected as treatment or control. Therefore:

$$(Y_{0i}, Y_{1i} \perp D_i \mid X) \text{ and } E(Y_{0i} \mid X_i, D_i=1) = E(Y_{0i} \mid X_i, D_i=0)$$

Performing the matching would be difficult due to the large set of observable variables, however, Rosenbaum and Rubin (1983) demonstrated how reducing the dimension to a univariate propensity score would be valid. In order to yield the conditional probability of being selected for treatment, I generate a propensity score based on observable variables such that: $P(X) = P(D=1 \mid X)$, and $0<P(X)<1$. The estimation of the propensity score requires the use of observable variables that are not influenced by the presence of the program; hence I choose baseline characteristics of the household that were present before the program started. It is also important to include a set of characteristics that identify the socioeconomic status of the household as well as any other relevant information available for the eligible adolescents.

After having a propensity score for each individual, and since it is true that $(Y_{0i}, Y_{1i} \perp D_i \mid P(X))$, I can rewrite equation (3) as:

$$ATT_i = E(Y_{1i} - Y_{0i} \mid D_i=1, P(X)) = E(Y_{1i} \mid D_i=1, P(X)) - E(Y_{0i} \mid D_i=1, P(X)) \quad (4)$$

Several different methods have been developed do the propensity score matching, some of them are:

1. Nearest neighbor matching: Select the first participant and find the non-participant with the closest score. That is, the non-participant with the value Pj that is closest to Pi.
2. Caliper matching: It's a variation of nearest neighbor matching which attempts to avoid bad matches. Define a region of tolerance (e.g. 0.01 to 0.001), and then randomly select the non-participant that most closely matches the propensity score of the participant.
3. Stratification matching: Create a set of intervals based on the propensity scores. Within each block there are r participants and n non-participants. Then use a bootstrap technique to match participants and non-participants.

The problem that arises for the evaluation of Avancemos using the propensity score matching technique is due to the existence of unobservable characteristics in selection which may lead to systematic differences between treated and untreated outcomes. The independence condition (also referred as conditional exogeneity assumption) defined by Rosenbaum and Rubin (1983) is not be satisfied if unobservables affect placement between both groups. Therefore, the unobservable characteristics would make the matching method unreliable. Heckman, Ichimura and Todd (1997) developed a difference-in-difference matching strategy to disentangle this selection problem.

Assuming that the unobservables are time invariant, the difference-in-difference method relaxes the conditional exogeneity assumption of single-difference estimators by using a baseline survey and a follow-up survey after the intervention. According to Todd (2008): "Difference-in-difference matching estimators identify treatment effects by comparing the change in outcomes for treated persons to the change in outcomes for matched, untreated persons. Difference-in-difference matching estimators allow for selection into the program to be based on unobserved time-invariant characteristics of individuals … difference-in-difference matching would solve the problem of the unobservables."

I Define time T=0 for the baseline period and T=1 for the post-intervention period, and by making use of the identification hypothesis, the difference-in-difference average treatment effect on the treated is:

$$ATT_i^* = [E(Y_{11i} \mid D_i=1, T=1, P(X)) - E(Y_{10i} \mid D_i=1, T=0, P(X))] -$$

$$[E(Y_{01i} \mid D_i=0, T=1, P(X)) - E(Y_{00i} \mid D_i=0, T=0, P(X))] \qquad (5)$$

In regression form, the difference-in-difference average treatment effect on the treated will be γ:

$$Y = \alpha + \beta D + \delta T + \gamma DT + \varepsilon \qquad (6)$$

Where:

α is the constant term,

β is the coefficient for the treatment dummy variable,

δ is the coefficient for the time variable,

γ is the coefficient for the treatment and time variables interaction term, and

ε is the error term.

VI Empirical Strategy

The treatment group consists of those individuals who reveal receiving the subsidy. I will use a logit model to calculate the propensity score of being selected based on 30 observable characteristics reported in the Costa Rican household survey that relate to regional, socio-economic and family characteristics. I assume that the selected observables are orthogonal to the treatment i.e. they help determine participation in the program, but that the program doesn't affect them. I also have to assume that those who didn't get the CCT aren't affected by this policy, i.e. they don't dropout from school only because they weren't selected. It is also important to recall that there's an externality that increases enrollment due to the subsidy, so I have to assume that higher class sizes do not lead to more dropouts or lower school quality.

For performing the propensity matching I use participation in Avancemos as the variable to be explained and the list of covariates chosen by Behrman et al. (2001) as explanatory variables: age, sex, total kids under age 12, total members of the household, ownership of the house, type of walls, type of floor, number of bedrooms, number of bathrooms, water availability, having a refrigerator, having a car, having a housing subsidy and having a job. Due to the extension of the Costa Rican Household Survey I'm able to include more observable variables that I consider relevant. Thus, I also include: regional characteristics (region of the country and a dummy for rural or urban zone), type of dwelling (house, apartment, etc), having electricity, a list of luxury goods (cellphone, computer, television, cable/satellite services, internet), type of school attended, the reason if dropping school, attending an informal education institution, relation to the head of the household and income. Since the variable income has several problems due to untrue responses, I use the head of the household's level of education (none, elementary, secondary or professional) as a proxy.

Concerning the effects of the covariates in the probability of selection in 2007, Table 6 shows the coefficients and Z statistics of the logistic regression. Most of the signs of the coefficients are as expected; lower income distinctiveness, and more schooling lead to a higher probability of selection into the program. In particular: total number of children under age 12 has a positive and significant effect on being selected, number of bathrooms has a negative and significant effect, having a car has a negative and significant effect and years of schooling of the head of the household has a negative and significant effect. For 2010 some of the durable goods coefficient's signs are reverted, perhaps due to the fact that the program has been already in place and that the selected households now have more income to spend in the consumption of durable goods while remaining selected. The number of bathrooms is still negative and significant and having a car remains negative and significant. Now total kids under age 12 has a negative and significant coefficient and total members a positive and significant coefficient. Having a refrigerator and TV now have positive and significant coefficients. Finally, head of the household's school level remains negative and significant. Full results are reported in Table 7.

After having propensity scores for all the eligible students and determining a region of common support I can move to create an appropriate control group from the untreated group. First I use stratification matching using Stata's command *atts*. I obtain that for the year 2010, the treated students have 0.678 more years of schooling than the control group. However, the existence of unobservable characteristics that affected selection into the program in 2007 makes this specification unreliable. To address this issue I will use other matching specifications that would allow me to incorporate a difference-in-differences strategy.

The next specification that I use is nearest neighbor matching. For each treated student, I find the untreated student with the closest propensity score. I do it in such a way in which there's replacement, so that the propensity scores are the closest even if the untreated student has been already selected before. I end up with a control group of the same size of the treatment group and as similar as possible based on observable characteristics. I do this process twice, once for the pre-intervention stage and again for the post-intervention stage.

For the pre-intervention stage I use the year 2007 because the program started nationwide in that year, and at the time the survey was made (June-July) the households would report being or not in the program while the program is still orthogonal to years completed (years are completed or failed by December of each year). I can't use 2006 as the baseline year because by then the program was only a pilot plan, and by the time the program started (May-June) the household surveys were already made; thus the program doesn't appear in the survey. For the post-intervention stage I use the most updated data possible. For this paper I use data from 2010, so the results will reflect the impact of the program's first 4 years, upcoming reviews would reflect the longer term effects.

It is expected that in the pre-intervention stage both groups would be almost identical and that in the post-intervention stage the effect of the program would be reflected in the treated group. The first comparison between groups is plotted in graph 7, and the second comparison in graph 8. Graph 7 shows how the groups were almost identical in years completed before the program, which is expected. Graph 8 shows how in 2010 the treated group has more years of schooling due to the participation in the program.

The last matching specification I use is caliper matching. I use the same methodology as in nearest neighbor matching, with the only difference that I exclude the *bad matches.* My upper bound for matches to keep is set to be 0.00000001, so I delete all the matches with greater propensity scores differentials than this cutoff. I consider that this last design will give the best estimates.

VII Results

The results are the solution to equation (6). Table 8 presents the results for difference-in-differences nearest neighbor matching and Table 9 presents the results for difference-in-differences caliper matching. The first column contemplates the whole sample, columns 2 and 3 are divided by sex, and columns 4 and 5 are divided by urban and rural zones. The coefficient of interest is the one on the interaction term, which is equal to the difference-in-difference average treatment effect on the treated.

With the nearest neighbor matching technique I get an increase in schooling of 0.339 years for the whole sample, which is significant at 5%. The boys only subsample (column 2) shows a notorious increase, 0.513 additional years of schooling, with 1% significance. All the other subsamples show positive but insignificant effects. The result are as expected, the program has increased years of school completed. The only potential problem that could affect these estimates is that *bad matches* are still included; caliper matching would take care of this issue.

Using caliper matching I obtain that years of school completed for the whole sample increased in 0.491 for the treated, and it's this result is significant at 1%. This means that after 4 years of the program, students in Avancemos stay in school for almost half a year more than if they weren't in the program. When splitting the sample to only boys I get that the additional schooling increases to 0.685 of a year and it's also significant at 1%. On the other hand, when splitting the sample to only girls I get a positive but insignificant effect of 0.321 years. For the urban subsample I get an increase in 0.705 years that is significant at 5%, and for the rural subsample I get and increase of 0.414 years of schooling that is also significant at 5%.

VIII Conclusion

CCT programs definitely affect the behavior of students. They could be an important tool for increasing school attendance and offset underage working, but they could also be very expensive programs. That's primarily why I considered very important to quantify the effects of the Avancemos program. According to my preferred specification, in the first four years of the Avancemos Program, selected students were able to increase their years of schooling in almost half a year. It seems that boys are the ones taking the bigger advantage of the program by increasing their schooling in more than two thirds of a year. Urban zones seem to have a slightly better use of the subsidy over rural zones.

Further research is needed to address very important issues that I left untouched in this paper. It is important to make reviews with newer data as it becomes available in order to quantify the long-term effects of the program, as well as trying to understand better why is the program increasing boys' years completed more than girls'. It is also relevant to follow studies such as Holzmann & Jorgensen (2000) in order to study the relations between intergenerational poorness and risk reduction, and see how this could be applied in the case of Costa Rica with the Avancemos Program. Another big topic left untouched in this paper is about returns from education, in future works I am interested in applying the Mincerian approach, so as to quantify the economic value of one more year of education and this way be able to make a cost-benefit analysis from the Avancemos long run results.

It certainly seems that this CCT program has allowed selected students to significantly increase their schooling. It is also true that the program is very expensive and that this money could have been spent differently (teacher incentives, school infrastructure, etc.), but what's important to recognize is that many underprivileged adolescents had the chance to stay in school a little longer, and that they now carry new potentials that only time could prove as a good or bad investment.

IX References:

- Behrman, Gallardo-Garcia, Parker, Todd and Velez-Granjalez (2006). "How conditional cash transfers impact schooling and working behaviors of children and youth in urban Mexico." Mimeo University of Pennsilvania.
- Behrman, Parker and Todd (2011). "Do conditional cash transfers for schooling generate lasting benefits". The Journal of Human Resources 46.1
- Behrman, Sengupta and Todd (2001). "Progressing through Progresa: An impact assessment of a school subsidy experiment." Economic Development and Cultural Change, 2005, vol. 54, issue 1, pages 237-75.
- Costa and De Oliveira (2008). "Avaliando Resultados de um Programa de Transferencia de Renda: O impacto de Bolsa-Escola sobre os Gastos das Familias Brasileiras".
- Duryea and Morrison (2004). "The Effect of Conditional Transfers on School Performance and Child Labor: Evidence from and ex-post Impact Evaluation in Costa Rica." Inter-American Development Bank, Research Department Working Papers; 505.
- Deaton (1997). "The Analysis of Household Surveys: A Microeconomic Approach to Development Policy." The Johns Hopkins University Press.
- Duflo (2001). "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment." American Economic Review 91 (4), 795-813.
- Garnier (2008). "Estadisticas Generales MEP 2008."
- Heckman, Ichimura and Todd (1998). "Matching as an Econometric Evaluation Estimator." The Review of Economic Studies. Vol. 65, No. 2 (Apr., 1998), pp. 261-294
- Instituto Nacional de Estadistica y Censos (INEC). "Encuesta Nacional de Hogares con Propositos Multiples 1999-2010."
- Mincer, J. (1974). "Schooling, Experience and Earnings". Columbia University Press.

- Molina & Fallas (2009). “Transferencias monetarias condicionadas en Costa Rica: el caso del programa Avancemos.”
- Parker, Rubalcava and Teruel (2008). “Evaluating Conditional Schooling and Health Programs.” Handbook of Development Economics, Volume 4. 2008 Elsevier.
- Ravallion and Wodon (1999). “Does Child Labor Dispace Schooling? Evidence on Behavioral Responses to an Enrollment Subsidy.” World Bank Policy Research Working Paper 2116.
- Rosenbaum, P. R. and Rubin, D. B. (1983). “The central role of the propensity score in observational studies for causal effects.” Biometrika 70 41--55.
- Schultz (2004). “School subsidies for the poor: evaluating the Mexican Progresa poverty program.” Workin Paper 834, Economic Growth Center, Yale University.
- Skoufias, Davis and Behrman (1999). “An evaluation of the selectiion of beneficiary households in the education, health, and nutrition program (PROGRESA) of Mexico.” IFPRI.
- Skoufias and Parker (2001). :Conditional Cash Transfers and Their Impacts on Child Labor and Schooling: Evidence from the Progresa Program in Mexico.” IFPRI
- Todd, P. (2008). “Evaluating Social Programs with Endogenous Program Placement and Selection of the Treated”. Handbook of Development Economics, Volume 4. 2008 Elsevier.
- Villatoro (2005). “Programas de transferencias monetarias condicionadas: Experiencias en America Latina.” Revista de la Cepal 86.

Appendix

Table 1 Characteristics of Students

Variable	2000	2005	2010
Fraction of Boys	.5007	0.0476	0.5061
Receiving Avancemos	--------	--------	0.324
In High School	0.6951	0.7501	0.8094
12 YO Mean Completion	5.341	5.840	5.940
13 YO Mean Completion	5.922	6.099	6.141
14 YO Mean Completion	6.321	6.680	6.779
15 YO Mean Completion	6.606	7.170	7.270
16 YO Mean Completion	6.989	7.583	7.802
17 YO Mean Completion	7.273	7.635	8.098
	100%	100%	100%
Observations	4156	4418	3904

Source: Costa Rican Household Survey

Table 2 Characteristics of Dropouts

Reason for Dropping School	2000	2005	2010
Has to Work	0.126	0.121	0.092
Rather Work*	------	0.058	0.061
Has to Work at Home	0.052	0.030	0.050
Can't Afford School	0.214	0.224	0.157
Access Problems to Schools	0.070	0.050	0.065
Problems Studying	0.114	0.134	0.129
Not Interested	0.322	0.320	0.396
Pregnancy or Marriage	0.042	0.038	0.030
Disability	0.040	0.025	0.026
	100%	100%	100%
*Not an option in the 2000 survey.			

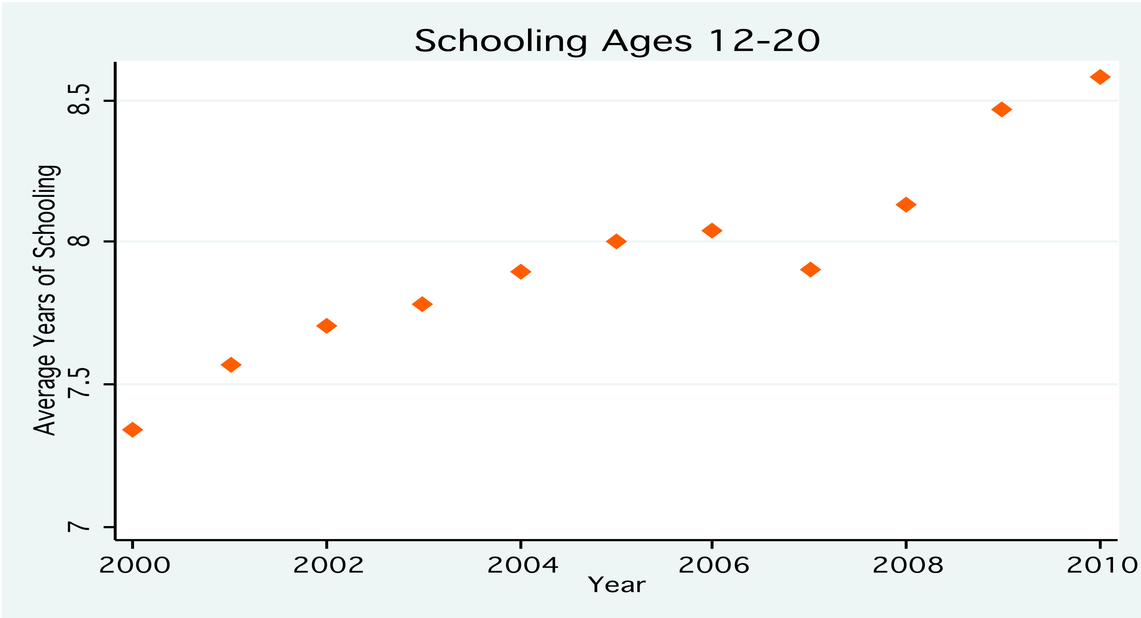
Source: Costa Rican Household Survey

Table 3 Head of the Households' behavior

Head of the HH of T=1	2007	2010
Mean Hours Worked	44.28329	44.36011
Standard Deviation	14.89754	14.36633
Observations	353	1058

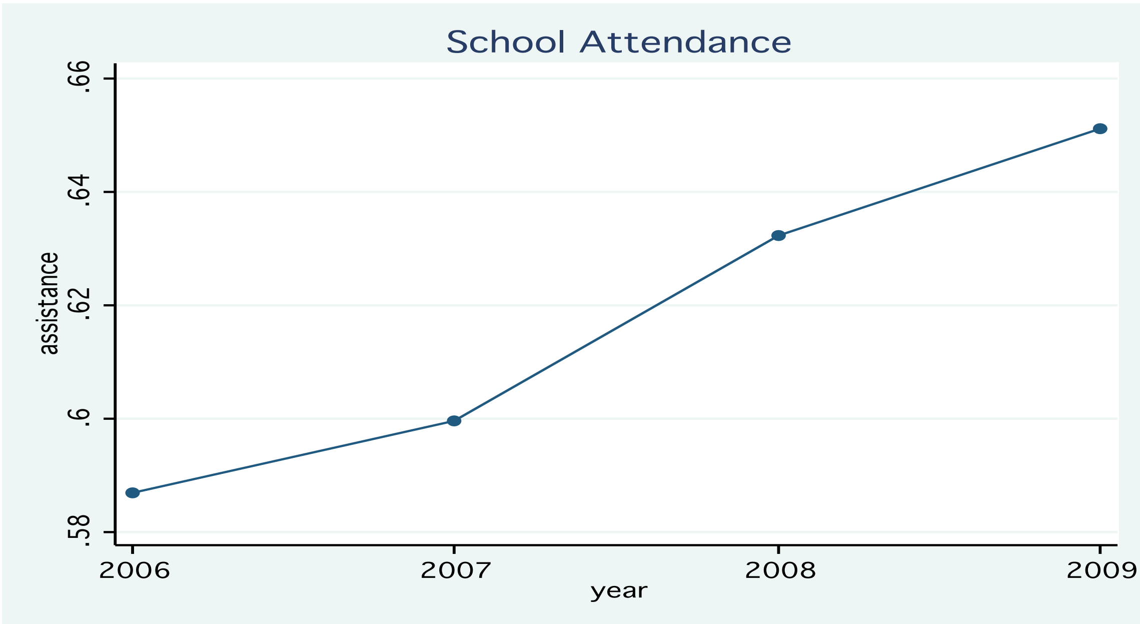
Source: Costa Rican Household Survey

Graph 1

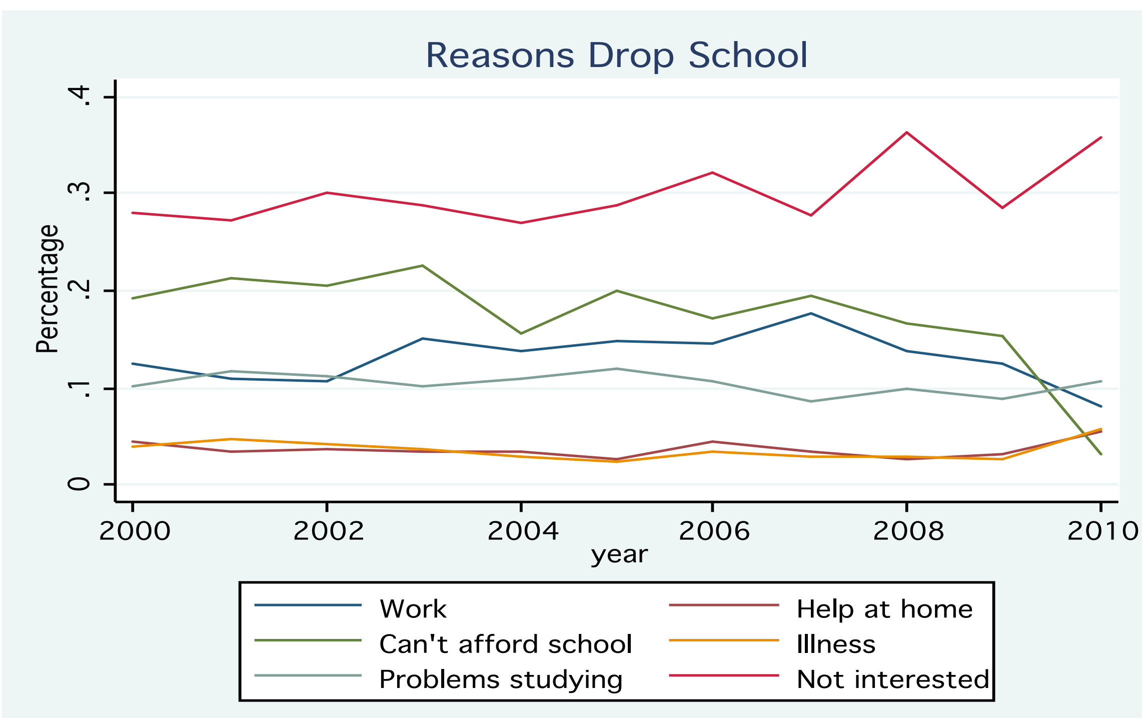


Graph 2

Graph 3

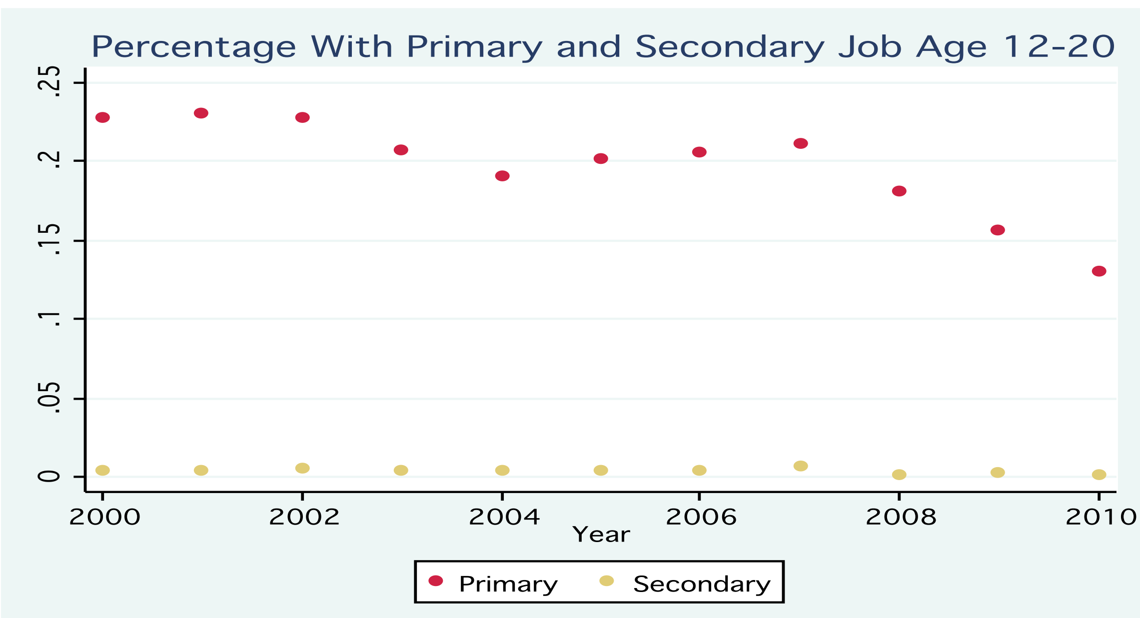


Graph 4

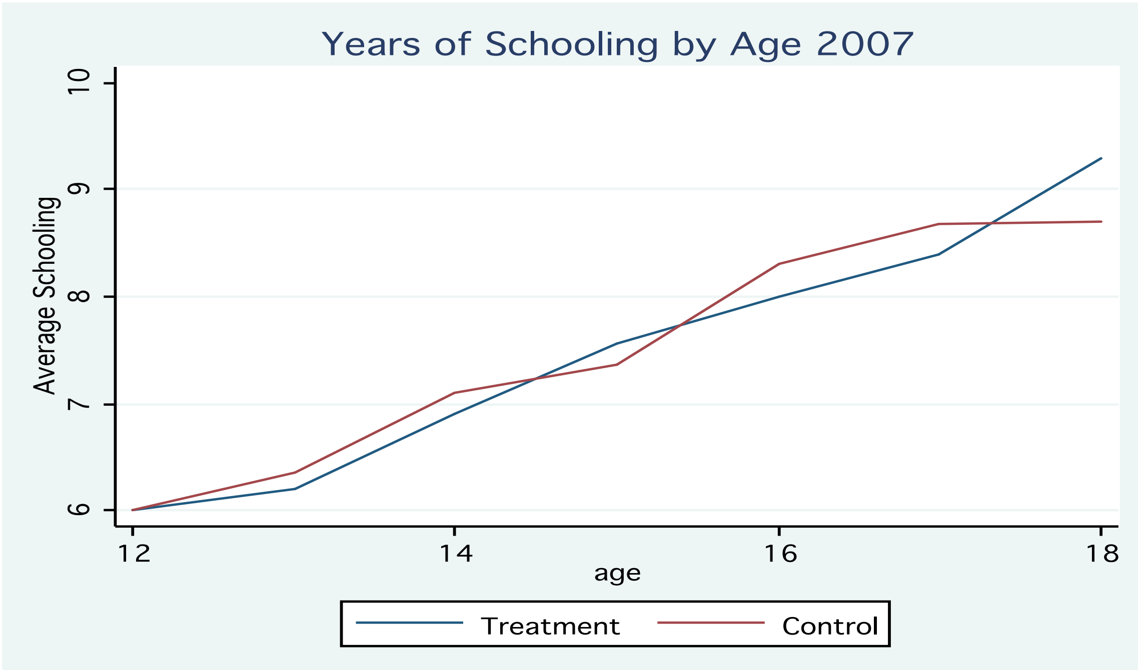


Graph 5

Graph 6



Graph 7



Graph 8

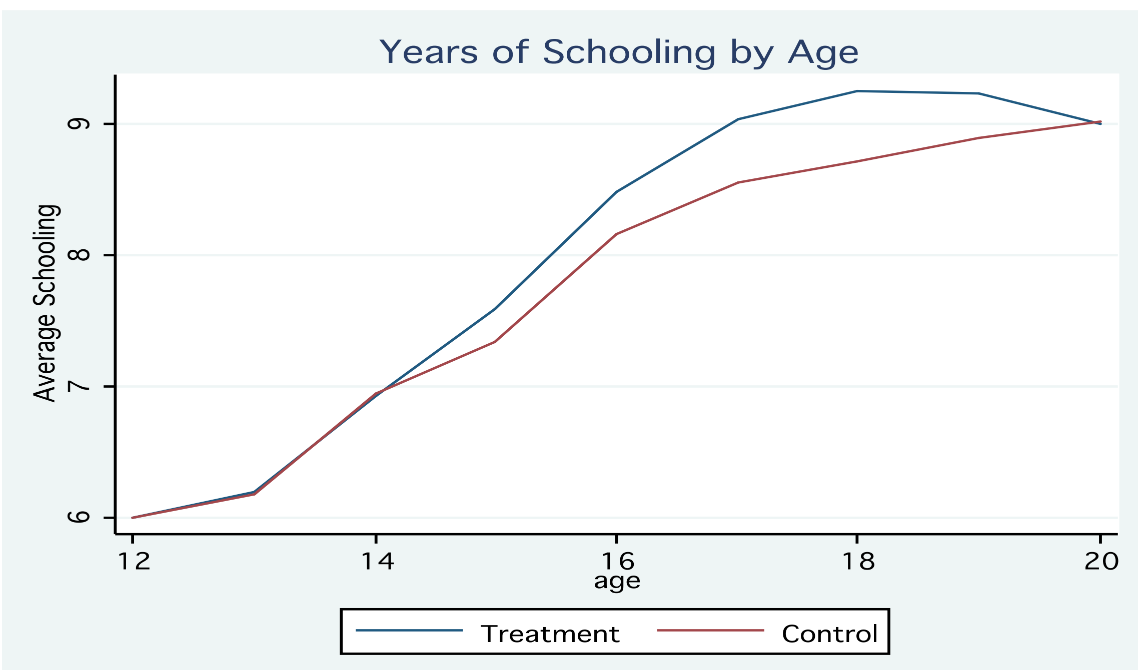


Table 4

Summary Statistics Eligible Students 2007

Variable	Obs	Mean	Std. Dev.	Min	Max
ReceivingCCT	6627	.0692621	.2539185	0	1
YearsCompl	6627	6.71254	2.060633	0	12
Age	6627	14.97148	2.024734	12	18
Sex	6627	1.47231	.4992704	1	2
Urban	6627	.3597405	.4799604	0	1
FamilySize	6627	5.153614	1.816001	1	15
TotalKids	6627	.9204768	1.037269	0	6
Cellphone	6627	.56134	.4962606	0	1
Refrigerator	6627	.9100649	.2861104	0	1
Computer	6627	.2796137	.4488433	0	1
Car	6627	.2716161	.4448266	0	1
TV	6627	.9435642	.2307787	0	1
CableTV	6627	.1982798	.3987341	0	1
SatelliteTV	6627	.0156934	.1242957	0	1
Internet	6627	.0746944	.2629175	0	1
HousingSubs	6627	.2194055	.413875	0	1
MaleHeadHH	6627	.7054474	.4558758	0	1
HeadHHSchool	6627	1.417383	.7524502	0	3
HastoWork	6627	.1404859	.3475167	0	1
HoursWorked	6627	5.128565	13.99344	0	60
Wage	6627	4177.048	17734.34	0	300000
FamilyIncome	6627	155349.8	260865.5	0	6680000

Source: Costa Rica Household Survey

Table 5

Summary Statistics Eligible Students 2010

Variable	Obs	Mean	Std. Dev.	Min	Max
ReceivingCCT	5825	.2401717	.4272243	0	1
YearsCompl	5825	7.054936	2.145002	0	12
Age	5825	15.05219	1.949312	12	18
Sex	5825	1.491502	.4999707	1	2
Urban	5825	.3860944	.4868945	0	1
FamilySize	5825	4.910901	1.699262	1	19
TotalKids	5825	.8109871	.9841548	0	8
Cellphone	5825	.7500429	.4330251	0	1
Refrigerator	5825	.9102146	.2858986	0	1
Computer	5825	.3848927	.4866117	0	1
Car	5825	.279485	.4487847	0	1
TV	5825	.9430043	.2318543	0	1
CableTV	5825	.2767382	.4474243	0	1
SatelliteTV	5825	.0248927	.1558115	0	1
Internet	5825	.1830043	.3867032	0	1
HousingSubs	5825	.2607725	.4390937	0	1
MaleHeadHH	5825	.6832618	.4652443	0	1
HeadHHSchool	5825	1.375107	.7469478	0	3
HastoWork	5825	.0727897	.2598133	0	1
HoursWorked	5825	2.457511	9.849971	0	60
Wage	5825	8387.603	39388.88	0	700000
FamilyIncome	5825	439844.8	508650.3	0	5360000

Source: Costa Rica Household Survey

Table 6

Probability of Selection for the Program 2007

Variable	Coefficient	Z Statistic
Age	0.057	(2.22)*
Sex	-0.055	(0.54)
FamilySize	-0.04	(1.01)
TotalKids	0.179	(2.82)**
TenureofHouse	0.039	(1.09)
Walls	0.038	(1.03)
Floor	0.093	(1.46)
Bedrooms	0.030	(0.55)
Bathrooms	-0.764	(3.97)**
Water	-0.042	(0.38)
TypeofHouse	-0.302	(2.61)**
Electricity	-0.059	(1.84)
Cellphone	-0.207	(1.89)
Refrigerator	0.497	(2.67)**
Computer	0.103	(0.64)
Car	-0.722	(4.32)**
TV	0.347	(1.55)
CableTV	-0.332	(1.74)
SatelliteTV	-0.880	(1.19)
Internet	-1.186	(1.96)
HousingSubs	0.184	(1.59)
Region	0.094	(3.12)**
Urban	0.342	(2.61)**
TypeofSchool	-0.938	(2.24)
HeadHHSchool	-0.347	(3.78)**
HastoWork	0.317	(1.01)
InformalEducation	-0.349	(1.76)**
RelationtoHeadHH	0.041	(0.92)
ReasonDrop	-1.961	(4.34)**
Constant	-2.570	(4.57)**
Observations	5694	

Absolute value of z statistics in parentheses

* Significant at 5%; ** Significant at 1%

Table 7

Probability of Selection for the Program 2010

Variable	Coefficient	Z Statistic
Age	0.121	(6.95)**
Sex	0.115	(1.76)
FamilySize	0.144	(5.20)**
TotalKids	-0.167	(3.68)**
TenureofHouse	-0.010	(0.41)
Walls	0.038	(1.60)
Floor	0.003	(0.06)
Bedrooms	0.007	(0.16)
Bathrooms	-0.545	(5.07)**
Water	-0.141	(1.14)
TypeofHouse	-0.057	(0.83)
Electricity	0.031	(1.14)
Cellphone	0.144	(1.86)
Refrigerator	0.391	(3.19)**
Computer	0.042	(0.44)
Car	-0.550	(5.98)**
TV	0.460	(2.90)**
CableTV	-0.417	(4.32)**
SatelliteTV	-0.234	(0.92)
Internet	-0.196	(1.54)
HousingSubs	0.482	(6.64)**
Region	0.038	(2.04)*
Urban	-0.120	(1.50)
TypeofSchool	-1.091	(4.06)
HeadHHSchool	-0.020	(0.35)
HastoWork	0.141	(0.49)
InformalEducation	0.103	(0.65)
RelationtoHeadHH	-0.065	(1.74)
ReasonDrop	-1.042	(7.86)**
Constant	-2.570	(4.57)**
Observations	5021	

Absolute value of z statistics in parentheses

* Significant at 5%; ** Significant at 1%

Table 8

	(1)	(2-B)	(3-G)	(4-U)	(5-R)
Years of Schooling					
Treatment					
	0.423	0.500	0.342	0.091	0.551
	(3.53)**	(2.93)**	(2.07)*	(0.42)	(3.86)**
Time					
	0.085	-0.023	0.160	-0.199	0.1.88
	(0.87)	(0.17)	(1.19)	(1.12)	(1.60)
Interaction (**ATT***)					
	0.339	**0.513**	**0.171**	**0.459**	**0.300**
	(2.45)*	(2.58)**	(0.91)	(1.83)	(1.81)
Constant					
	6.787	6.541	7.054	7.231	6.618
	(80.06)**	(54.17)**	(60.44)**	(46.87)**	(65.44)**
Observations					
	3380	1659	1721	982	2398

Absolute value of t statistics in parentheses

* Significant at 5%

** Significant at 1%

Table 9

	(1)	(2-B)	(3-G)	(4-U)	(5-R)
Years of Schooling					
Treatment					
	0.347	0.473	0.204	-0.063	0.536
	(2.48)*	(2.37)*	(1.06)	(0.26)	(3.17)**
Time					
	-0.175	-0.267	-0.111	-0.510	-0.051
	(1.44)	(1.53)	(0.67)	(2.49)*	(0.34)
Interaction (**ATT***)					
	0.491	**0.685**	**0.321**	**0.705**	**0.414**
	(2.86)**	(2.78)**	(1.37)	(2.40)*	(1.97)*
Constant					
	6.730	6.474	7.007	7.290	6.478
	(68.07)**	(46.01)**	(51.43)**	(44.60)**	(53.82)**
Observations					
	1786	909	877	601	1185

Absolute value of t statistics in parentheses

* Significant at 5%

** Significant at 1%

Stata Codes:

*/ Summary Statistics

```
*/ Creating a Family ID:

gen con= consecu*100

gen familyid=con+ vivienda

sort familyid, stable

*/ Creating a Dummy Variable for Male Head of HH:

bysort familyid: gen head = (_n == 1)

gen mhead=1 if sex==1 & head==1

bysort familyid: egen malehead = max(mhead)

replace malehead=0 if malehead==.

*/ Creating a variable for Head of HH Schooling

gen schyrs=.

replace schyrs = 0 if yearconc==0

replace schyrs = 1 if yearconc>0 & yearconc<20

replace schyrs = 2 if yearconc>19 & yearconc<40

replace schyrs = 3 if yearconc>39

gen hsch = schyrs if head==1

bysort familyid: egen hhhsch = max(hsch)

*/ Creating Income per Family and Dummies for Durable Goods

gen totincome=incomejob

bysort familyid: egen familyincome = sum(totincome)

gen fincomepc = familyincome/total

drop if age<12 | age>18

drop if scholarsh!=0 & scholarsh!=1

replace cel=0 if cel==2

replace fridge=0 if fridge==2

replace computer=0 if computer==2

replace dvd=0 if dvd==2

replace car=0 if car==2
```

```
replace tv=0 if tv==2

replace cable=0 if cable==2

replace satellite=0 if satellite==2

replace internet=0 if internet==2

replace hsub=0 if hsub==2

drop if hsub>1

*/ Normalizing Years of Schooling from 1 to 12

replace yearconc=1 if yearconc==11

replace yearconc=2 if yearconc==12

replace yearconc=3 if yearconc==13

replace yearconc=4 if yearconc==14

replace yearconc=5 if yearconc==15

replace yearconc=6 if yearconc==16

replace yearconc=7 if yearconc==21 | yearconc==31

replace yearconc=8 if yearconc==22 | yearconc==32

replace yearconc=9 if yearconc==23 | yearconc==33

replace yearconc=10 if yearconc==24 | yearconc==34

replace yearconc=11 if yearconc==25 | yearconc==35

replace yearconc=12 if yearconc==26 | yearconc==36

drop if yearconc>12

replace job=0 if job==2

*/ Labeling

rename scholarsh ReceivingCCT

rename zone Urban

replace Urban=0 if Urban==2

rename totua TotalKids

rename total FamilySize

rename cel Cellphone

rename fridge Refrigerator

rename computer Computer

rename car Car

rename tv TV

rename cable CableTV
```

```
rename satellite SatelliteTV

rename internet Internet

rename hsub HousingSubs

rename sex Sex

rename age Age

rename yearconc YearsCompl

rename job HastoWork

rename hoursjob HoursWorked

rename incomejob Wage

rename familyincome FamilyIncome

rename fincomepc FamilyIncomePC

rename malehead MaleHeadHH

rename hhhsch HeadHHSchool

sum ReceivingCCT YearsCompl Age Sex Urban FamilySize TotalKids Cellphone Refrigerator Computer Car TV CableTV
SatelliteTV Internet HousingSubs MaleHeadHH HeadHHSchool HastoWork HoursWorked Wage FamilyIncome, separator(1)

*/ Head of the Household working hours before and after program:

gen con= consecu*100

gen familyid=con+ vivienda

sort familyid, stable

bysort familyid: gen head = (_n == 1)

gen hjobhours=hoursjob if head==1

replace hjobhours=100 if hjobhours==0

replace hjobhours=0 if hjobhours==.

bysort familyid: egen hhhours=sum(hjobhours)

drop if hhhours>70

drop if scholarsh!=1

sum hhhours

*/ Reasons for dropping out

collapse (mean) assistance, by(year)

twoway connected assistance year

forvalue i = 1(1)10 {

 gen reason`i' = 1 if reasondrp == `i'

 replace reason`i' = 0 if reason`i' == .

 }
```

```
collapse reason1-reason10 if reasondrp > 0, by(year)

tab  reasondrp if reasondrp !=0

twoway line reason1 year || line reason2 year || line reason3 year || line reason4 year
```

***/ Propensity Score:**

```
pscore scholarsh relation InfEduc reasondrp Age Sex FamilySize TotalKids TenureH Walls Floor Bedrooms Bathrooms Water Refrigerator Car HousingSubs HastoWork Region Urban TypeH Elect Cellphone Computer TV CableTV SatelliteTV Internet HeadHHSchool typesch, pscore(pscore) blockid(block) logit
```

***/ Stratification Matching:**

```
atts yearconc scholarsh, pscore(pscore) blockid(block)
```

***/ Creation of the Control Group:**

```
egen pid=seq()

sort avancemos pid

gen pidcontrol=.

gen mindist=.

quietly forval i=2475/`=_N' {

local mindist=1

 forval j=1/2474 {

  local dist=(pscore[`j']-pscore[`i'])*(pscore[`j']-pscore[`i'])

  if `dist'< `mindist' {

   replace pidcontrol=pid[`j'] in `i'

   replace mindist=`dist' in `i'

   local mindist=`dist'

  }

 }

}

preserve

keep if avancemos==0

drop pidcontrol mindist pscore

renvars, postfix(_c)
```

```
rename pid_c pidcontrol

sort pidcontrol

tempfile controls

save `controls', replace

restore

keep if avancemos==1

sort pidcontrol

merge pidcontrol using `controls'

tab _merge

drop if _merge==2

drop _merge

drop if mindist>.00000001
```