Alameda 3363
Estación Central-Santiago
Tel. +56 2 7180765
www.usach.cl

Universidad de Santiago de Chile

Departamento de Economía

Working Paper Series

Did PROGRESA send dropouts back to school?

Autor: María Nieves Valdés

WP 2011 - Nº 09

Did PROGRESA send dropouts back to school?

María Nieves Valdés*†
Universidad de Santiago de Chile

Abstract

This paper analyzes the effect of PROGRESA education grants on school enrollment. It looks at its effect on total school enrollment and in particular on school enrollment of dropouts, i.e. those children who face a re-enrollment decision since they were not enrolled in school the year prior to the implementation of the PROGRESA program. Estimates of the impact of PROGRESA education grants on dropouts and non-dropouts are obtained applying difference estimation and maximum likelihood estimation of a reduced form equation for schooling decision. Differences in results between both groups of children are discussed looking at the distribution of marginal effects. PROGRESA did send dropouts back to school. It had a larger effect on dropouts than on non-dropouts. However, for the particular group of girls who dropped out of school just before attending secondary school PROGRESA grants only had a minor effect. This last finding highlights the fact that determinants of the schooling decision are different for young girls and that PROGRESA grants do not provide a strong enough incentive to send them back to school.

JEL classification C21; C23; I28; I38

Keywords Anti-poverty program evaluation; School enrollment; Re-enrollment decision; Heterogeneous program effects; Correlated random effects model

*Email: nieves.valdes@usach.cl.

†Special thanks to Pedro Albarrán and Ricardo Mora for their advice, remarks and comments. I am grateful to Cesar Alonso, Jesús Carro, Ramiro de Elejalde, Sara de la Rica, Rodolfo Stucchi, Alejandra Traferri, and Marcos Vera for insightful comments on preliminary drafts of this paper. I also acknowledge comments by seminar participants at Universidad Carlos III de Madrid, Universidad de Oviedo, EEA-ESEM 2006 Meeting, and ESPE 2007 Meeting.

1 Introduction

The key role of education as an anti-poverty and pro-growth policy via the effect of education on the accumulation of human capital is widely recognized by both policy makers and the economic literature. Not surprisingly, policies aimed at increasing education have received a great deal of attention over the last 50 years.

Government policies can influence educational choices by affecting the three main factors that determine school enrollment and completion: demand, supply, and educational policy. On the demand side, policy makers have decreased direct costs by reducing or eliminating tuition or transport costs. Thet have also implemented cash transfers programs aimed at decreasing opportunity costs of wage income and/or home production forgone. Young people may also leave school for supply side reasons such as the availability or quality of education. Increasing the availability of primary and secondary schools and improving the quality of the education by reducing class size and strengthen the qualifications of teachers, have long been included in educational policy portfolios. On the institutional side, governments have been increasing the number of years of compulsory education. In most countries children have to attend school until the age of 16 years old, and many countries are imposing compulsory education until the secondary school level.

Over the last two decades there has been a widespread use of programs aimed at fostering the accumulation of human capital in the developing world. The increasing use of anti-poverty programs has been accompanied by comprehensive evaluations about the actual effectiveness of these programs. This has occurred not only because of the intrinsic interest in these programs but also because anti-poverty programs represent an important financial effort, both by governments and international institutions which often provide additional funds.

This paper focuses on a Mexican anti-poverty program for rural communities, called PROGRESA (Education, Health and Nutrition program), first implemented in 1997 by the Mexican Federal Government. The program comprise three major areas one of which, the subject of this paper, is education. In particular, program beneficiaries are given financial aid conditional on school attendance. This paper analyzes the effects of such grants on school enrollment for two different groups of beneficiaries, dropouts and non-dropouts. The identification of these effects relies on the randomized assignment of the program benefits.

The evaluation of the program was conducted by the Mexican Federal Government and by external local and international evaluators such as the National Institute of Public Health (INSP, Mexico), Research and Advanced Studies Center in Social Anthropology (CIESAS, Mexico), International Food Policy Research Institute (IFPRI), and Research and Educational Documentation Center (CIDE, Spain)[1]. The evaluation efforts have resulted in an extensive literature by authors like Orazio Attanasio, Jere R. Behrman, David Coady, Costas Meghir, T. Paul Schultz, Emmanuel Skoufias, Petra Todd and Kenneth Wolpin, among others. In this literature authors estimate average treatment effects (ATE) exploiting the randomized assignment of the program. Their results prove the success of PROGRESA in increasing enrollment rates for those children who received the grants, and they agree in that this positive effect is higher on girls and on children who attend secondary school.

Although the ATE can be a good general characterization of the overall (average) effect, it is obvious that any program will have a different impact on different individuals. Some of them will benefit a lot whereas others will not. Therefore, it is also important to take into account individual heterogeneity.

The main contribution of this paper is to track the differential effect of the program on individuals that dropped out of school before the program started. These children are facing a re-enrollment decision that may imply higher direct and/or indirect costs of schooling than the costs faced by the average child. Moreover, dropouts are different from the average child in some observable characteristics related to the schooling decision. Thus, we can expect a different effect of the program on them. The methodologies applied are difference estimation and maximum likelihood estimation of a reduced form equation of education choice. For both cases the randomized design of PROGRESA is exploited. The outcome is the causal effect of the education component of this program.

The estimated marginal effects provide evidence on the existence of differences between the impact of PROGRESA grants on the overall target population and their impact on the children who face a re-enrollment decision. The rise in boys school attendance caused by PROGRESA grants is higher for dropout boys in both levels of education, primary and secondary school. However, for girls in secondary school who dropped out in 1997 or before, the grant has a negligible effect on their decision to reentry in school. Among dropouts in secondary school the impact of PROGRESA grants is lower for girls even though they receive more money than boys.

The structure of the paper is as follow. Section 2 presents the main features of

[1]Corresponding web pages: Federal Government: www.oportunidades.gob.mx; INSP, www.insp.mx; CIESAS, www.ciesas.edu.mx; IFPRI, www.ifpri.org; CIDE, www.mec.es/cide/

the program and a brief review of the literature evaluating PROGRESA. Section 3 discuses factors that influence the enrollment decision and presents differences to the re-enrollment decision made by dropouts. Section 4 describes characteristics of the PROGRESA data base. It provides some main statistics that focus on the differences between dropouts and non-dropouts. In Section 5 results for the difference estimation of the effects of the program are presented for both groups and are analyzed separately. Section 6 introduces a reduced form equation for the schooling decision including PROGRESA education grants variables. Section 7 presents maximum likelihood estimates of a probit model for schooling decision, comparing results for non-dropouts and dropouts. Finally, Section 8 concludes the paper with its main results and some suggestions for future research.

2 The PROGRESA program and its education component

The Education, Health and Nutrition program, PROGRESA, was implemented by the Federal Government of Mexico in 1997, with the aim of helping the poorest families in rural communities. A fundamental characteristic of the program is that aid is conditioned on a specific behavior of the beneficiary. This conditionality tries to guarantee that the program does not lead to undesired outcomes, such as distortions in work decisions, and that it successfully accomplishes its initial objectives.

The program comprises actions in three major areas: education, health and nutrition. The education component includes monthly grants for children of a family qualified as beneficiary. They need to be less than 18 years old, enrolled in school between the 3^{rd} year of primary school and the 3^{rd} year of junior secondary school, and to fulfill a minimum attendance requirement. The grants are not based on academic achievement. A child who does not pass a grade is still eligible for the grant in the following year. But if the child fails the same grade twice, she/he losses eligibility. The grant increases by years of schooling. In the junior secondary level the grant is slightly higher for girls, since there exist evidence that in poor families girls are more likely to dropout of school and that they also tend to dropout earlier than boys. Additionally, beneficiaries receive an annual grant for school supplies. The health component of the PROGRESA program consists of a basic package of free health services, nutritional supplements, and informative talks on health, nutrition, fertility, and hygiene. Special attention is paid to pregnant women and children younger than five years. Finally, the nutrition component of the program supplies beneficiary families with a monthly monetary payment intended to improve amount and diversity of food consumption and thus increase the nutritional status, in particular of children. This aid is independent of residence, and size, and composition of the family. All aid is given to the mother of the family as there exist evidence that mothers are better than fathers at allocating family resources.

A family is qualified as being poor and thus eligible for the program according to a single index. This index contains information on family income and housing like presence of running water, etc.[2]

Some numbers can provide a better understanding of the extent and significance of PROGRESA as an anti-poverty policy. In 1997 the program reached 6,357 communities, giving aid to 300,705 families. This implied transfers of 34 million USD (approx. 340 million Mexican pesos). After two years of being implemented the program included nearly 2.6 million families in 72,345 communities in all 31 Mexican states. It reached around 40% of all rural families and nearly 12% of all families in Mexico. Total annual transfers of the program in 1999 were around 710 million USD, equivalent to 0.15% of Mexican GDP. 40% were educational transfers, 42% food transfers and 18% was spent on health transfers. Among the total annual cash transfers of 578 million USD, food transfer accounted for 49%. The remaining 51% went to education. In 1999 the program distributed 273 million USD in education grants[3].

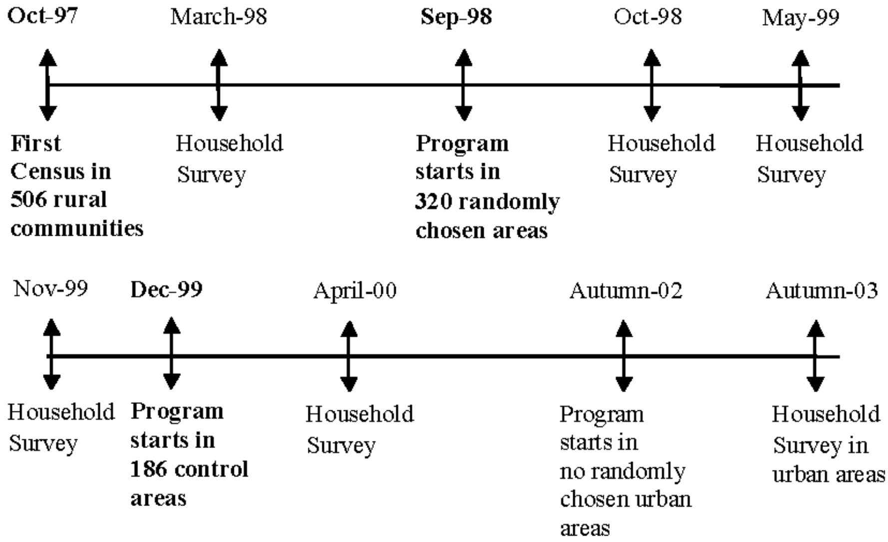
Given the financial importance of PROGRESA, Mexican authorities have intended to evaluate the program since its beginning, not only to measure results and impacts but also to provide information that allow for a redesign of policies. Accordingly, in 1997 and 1998 a high quality data set was collected in 506 communities where the program was to be implemented, and several surveys were carried out afterwards. In October 1998, the program was implemented in 320 randomly selected communities (treated communities) while in the remaining 186 communities (control communities) the implementation was postponed until December 1999[4]. In Figure 1 below, I present the timing of the program.

[2]For a complete analysis of the targeting see Skoufias, Davis, and Behrman (1999a) and Skoufias, Davis, and Behrman (1999b).

[3]For more details on PROGRESA costs see Coady (2000).

[4]The quality of the randomization has been extensively documented in Behrman and Todd (1999), who conclude that, at least at community level, the implementation of the random assignment was performed successfully.

Figure 1: Timing of the PROGRESA program



The evaluation of the program was conducted by the Mexican Federal Government and by external local and international evaluators such as the National Institute of Public Health (INSP, Mexico), Research and Advanced Studies Center in Social Anthropology (CIESAS, Mexico), International Food Policy Research Institute (IFPRI), and Research and Educational Documentation Center (CIDE, Spain). The evaluation efforts have resulted in an extensive literature by authors like Orazio Attanasio, Jere R. Behrman, David Coady, Costas Meghir, T. Paul Schultz, Emmanuel Skoufias, Petra Todd and Kenneth Wolpin, among others.

This paper is closely related to Schultz (2004). In his paper Schultz presents an extensive evaluation of the education component of PROGRESA. The author performs pre-program comparisons to check the randomization of the design, and he calculates difference and difference-in-difference estimators by gender and grade which allow him to quantify the program's causal effect. To validate difference estimations he shows results of maximum likelihood estimation of a reduced form equation of the school enrollment decision. He concludes that the program has effectively reached its goal since he finds positive and large post-program differences in enrollment rates of comparably poor children in treatment and control communities.

Other related relevant papers are Attanasio, Meghir, and Santiago (2005) and Todd and Wolpin (2003), who follow a structural approach to evaluate PROGRESA. They can thus simulate the effects of counterfactual programs and they can identify alternative subsidy schemes with a greater impact on schooling decisions.

Attanasio, Meghir, and Santiago (2005) estimate the structural parameters of a standard model of education choices that considers schooling as an individual decision. Similarly to Schultz (2004), they find that PROGRESA has a positive effect on the school enrollment of children, especially after the completion of primary school. They also show that a revenue neutral change in the program that increased the grants for junior secondary school children while eliminating the ones for primary school children, would have a substantially larger effect on enrollment of secondary school children, while having only minor effects on the enrollment of primary school children.

Todd and Wolpin (2003) estimate a dynamic behavioral model of parental decisions about fertility and children schooling. Their paper differs from Attanasio, et al in two main aspects. Todd and Wolpin (2003) model schooling as a family decision. They use data from the control group prior to the experiment in the validation and estimation of the model, i.e. they use only pre-program information in the estimation of the parameters of interest. They then apply the model to analyze the effectiveness of alternative policies to increase enrollment rates.

3 Enrollment vs. re-enrollment decision

3.1 Influential factors for the enrollment decision

From an economic point of view, the school enrollment decision is taken based on the private price of schooling. The total price of schooling includes tuition fees, direct costs of attending school, such as clothing, books, materials and transportation costs but also the opportunity cost if attending school. Since in Mexican rural communities public schools are mostly tuition-free, the main component of the price of schooling is the opportunity cost of time. A student could devote her/his time spent at school to other activities, such as paid work, farming, or any other productive activity at home. PROGRESA directly reduces the price of schooling through grants and aid for school supplies. From this reduction in the price of schooling, we would expect a positive effect of the program on enrollment rates.

The main component of the opportunity cost of schooling is the rural wage a child can earn as farming or home production activities are difficult to measure in monetary terms. Unfortunately only a small fraction of communities report such information. As a proxy we consider the agricultural wage for adult male.

In communities with high salaries we expect that children are less likely to go to school, because they face a higher opportunity cost. Additionally, medium and large cities have more developed labor markets that usually offer higher wages. So we expect a child residing near a metropolitan area or near the main city of her/his municipality to be more likely to dropout of school and to work instead.

Transportation cost are an important direct cost of schooling for children attending junior secondary school. Only in 25% of all communities under study have a proper junior secondary school. A reasonable proxy for this cost is the distance from the community where the child resides to the nearest one that has a secondary school. In all the communities studied there exist at least one primary school, so we can set transportation costs for primary school children equal to zero.

Given that a child's schooling is a family decision, it is necessary to analyze family characteristics that may influence this decision. There is a general agreement that more educated parents are more likely to send their children to school. If the father lives at home and works we expect his children to be more likely to go to school, as the financial situation of the family is more stable. Health and work status of the head of the household are also relevant for a child's schooling decision. If the head of the household was unemployed or ill for some weeks before the children should have been enrolled at school it is likely that the children are sent to work.

3.2 The re-enrollment decision

The focus of this paper is not the overall population but those children that are making a re-enrollment choice, i.e. dropouts. Drop-outs are those who have made the decision of not attending school at some point in time and were not enrolled in October 1997, before the implementation of PROGRESA. These children were not receiving enough incentives to go to school, to improve their educational level, and to contribute finally to human capital accumulation and the development of their communities. Regarding the aim of PROGRESA, dropouts are thus an important target of the program.

Is there any reason to think that PROGRESA education grants could have a different impact on enrollment rates of dropouts than on those of non-dropouts? A hypothetical answer to this question can be made based on observable differences between both groups. In particular, we can look at information provided by the pre-program census (October 1997) and interpret it referring to the conclusions from the previous subsection. Some numbers are given in Table 1 below.

As I expect values of some variables to be different for different levels of schooling and gender, data is presented separately for primary and secondary school children and for girls and boys. The primary school sample includes all children aged 6 to 18 who have completed 0-5 years of schooling and are thus qualified to enroll in primary school grades 1 to 6. In the secondary school sample I considered all children aged 11 to 18 who have completed at least 6 years of schooling and are thus qualified to enroll in junior or senior secondary school.

Table 1: Difference in averages and proportions of selected variables between Dropouts and Non-dropouts (pre-program census)

Variable name	Primary		Secondary	
	Female	Male	Female	Male
Percentage of children belonging to a poor family	4.2***	3.7***	8.0***	3.7*
Mother's schooling (years)	-1.5***	-1.3***	-1.0***	-1.0***
Percentage of children with father living at home	-3.4**	-5.8***	0.9	-0.1
Number of siblings enrolled at school	-1.7***	-1.6***	-1.2***	-1.3***
Distance to secondary school (km)	0.92***	0.64***	1.11***	0.74***
Distance to metropolitan area (km)	-17.6***	-19.3***	-22.8***	-31.4***

* Statistical significance = 10%. ** Statistical significance = 5%. *** Statistical significance = 1%.
The statistical significance of the differences is tested using tests for equality of means and proportions.

Having a more educated mother increases the probability of enrollment for non-dropouts. The positive effect of parents' education on the education of their children is well known. Additionally, the proportion of children of families with the father living at home is higher, making them more likely to attend school (at least in primary school). Non-dropouts face lower direct cost of schooling since they live closer to a secondary school.

On the other hand, dropouts reside closer to metropolitan areas and the main cities of their municipalities. Drop-outs reside in communities with higher wages. These facts imply a higher opportunity cost of schooling leading to a lower probability of re-enrollment for dropouts. Moreover, a higher proportion of them belong to a poor family, making them more likely to work and not to attend school.

Additional information is given in the surveys that were carried out after PROGRESA started. In particular, these surveys ask why the child was not enrolled in school. For dropouts the answers are: "There was not enough money" (47%), "She/he did not like going to school" (26%), "The school was very far away" (9%), and "Her/his help was needed at work or at home" (4%). Clearly the main reason for not attending school are financial restrictions at home making the alternative of working even more attractive.

Summing up, there exists enough evidence to conclude that dropouts have more incentives to work rather than to attend school compered to non-dropouts. Given the higher opportunity cost of schooling that dropouts face and since the grants are a monetary incentive[5], I expect the effect of the program on dropouts to be larger than on the average child. Thus, the proposed hypothesis based on observable characteristics is that the program has a stronger effect on the schooling decision of dropouts. However, we do not know in which direction any unobservable characteristic of the child, like ability or ambition, could affect the schooling decision and if it could affect the schooling decision of dropouts and non-dropouts in a different way.

4 Data base and descriptive statistics

Since in this paper I analyze the grants' impact conditional on the schooling decision children have made before the implementation of the program, only post-program information can be used. From the education component of the PROGRESA post-program surveys (October 1998, May 1999 and November 1999) a matched panel sample for children aged 6 to 18 can be obtained. This panel includes 74,427 observations, 45,666 (61%) in the treatment group (individuals residing in a community where PROGRESA grants were implemented in September 1998) and 28,761 (39%) in the control group.

Before going into detail on the description of the data base, three comments should be made. First, there exists a maximum amount of aid a household can receive by means of the education component of the program. Those maximum amounts are updated every six month (as it happens with grants). When the maximum is reached each child receives only a percentage of the grant. Unfortunately, the exact amount each child receives is not reported in the data base. For this reason what can be measured is only the effect on school enrollment of the "potential grant". Using this measure for the effect of the program we may overestimate the "actual grant" effect. If the child's family is not receiving the maximum amount the potential grant

[5]The monthly agricultural wage is around $ 500, but a child actually earns less than this amount. A secondary school child's grant is approximately $ 250. These numbers show how important PROGRESA grants are as an additional source of family income.

coincides with the actual grant. In the PROGRESA data base the average number of children in a family is 4. This makes it very likely for a family to attain the maximum amount of aid.

The second comment is about the treatment group. Around 5% of those children fulfilling the requirements to obtain the grant are not receiving it[6]. The reason for this is not available in the data base. The grant amount for them is set to zero.

Finally, the variable reflecting the stock of education (years of schooling completed) presents some inconsistencies along the waves of the surveys. 29% of the observations show some of these inconsistences. For this reason I perform a hand-correction of the variable "stock of education" using the information available in five waves, that cover information on four academic years. For all observations to be corrected I have at least two consistent combinations of enrollment-age-schooling, and I correct the remaining points making it compatible with that sequence. In all cases I preserve the information related with enrollment to school and the age. Given the exogenous variation generated by the random assignment of the grant, there is no reason to think that the measurement error in the stock of education will bias the estimation of the grant effects. Nevertheless, it may induce some bias in the estimation of average treatment effects by level of education.

In terms of the data base, dropouts are those individuals aged 8 to 18 in the post-program surveys who were not enrolled at school in the first census (October 1997). The re-enrollment or dropouts panel includes 6,948 observations, 4,155 (60%) in the treatment communities and 2,793 (40%) in the control communities.

Table 2 presents a set of descriptive statistics that characterize the population in the dropout panel.

[6]The exact numbers are 5.62% for non-dropouts and 5.14% for dropouts.

Table 2: Descriptive statistics for Drop-outs (post-program surveys)

Variable	Primary		Secondary	
	Female	Male	Female	Male
Sample size	1,310	1,490	2,431	1,717
Enrollment rate	0.526	0.528	0.258	0.259
Percentage of treatment communities	57.5	59.9	59.2	62.3
Percentage of children belonging to a poor family	92.9	92.9	84.2	83.1
Percentage of children eligible for receiving a grant	26.9	31.8	44.5	46.7
Grant (for grant different from zero) (pesos)	118.7	122.7	250.7	237.6
	(31.4)	(30.5)	(22.9)	(19.2)
Mother's schooling (years)	1.5	1.6	2.1	2.0
	(2.1)	(2.0)	(1.9)	(2.1)
Percentage of children with head of household ill	6.3	6.5	6.8	6.9
Percentage of children with head of household employed	89.2	89.5	90.2	87.5
Percentage of children with father not living at home	13.3	16.2	9.6	10.9
Number of girls from 5 to 16	2.0	1.0	1.9	1.0
	(1.2)	(1.0)	(1.2)	(1.0)
Number of boys from 5 to 16	1.0	1.9	1.0	1.9
	(1.0)	(1.3)	(1.0)	(1.2)
Number of children under 5	0.9	0.9	0.6	0.5
	(1.0)	(1.0)	(0.9)	(0.9)
Number of adult women	1.7	1.6	2.0	1.8
	(0.9)	(0.8)	(1.0)	(1.0)
Number of adult men	1.6	1.8	1.8	2.2
	(1.0)	(1.0)	(1.4)	(1.1)
Number of siblings enrolled at school	2.1	2.3	2.3	2.0
	(1.5)	(1.4)	(1.4)	(1.5)
Distance to secondary school (km)	3.2	3.0	2.8	2.6
	(3.3)	(3.0)	(2.0)	(1.9)
Percentage of children that have a secondary school in their community	18.0	21.9	15.1	18.5
Distance to nearest metropolitan area (km)	131.8	131.0	129.2	125.2
	(64.8)	(59.7)	(73.2)	(66.4)
Distance to the main city of her/his municipality (km)	11.2	11.2	12.0	10.7
	(7.2)	(7.2)	(9.4)	(7.4)
Community daily agricultural wage (pesos)	29.4	30.2	33.4	31.8
	(11.4)	(11.2)	(11.9)	(11.6)

Standard deviations are in parenthesis *(continued in Appendix as Table 1-continued)*

Table 6 of the Appendix reports similar statistics for non-dropouts.

In both tables variables are somewhat different for girls and boys, as expected. Also, we see different variable values for children in primary and in secondary school. Hence, I want to use an estimation strategy that will allow for differences in the program's effects by gender and by level of education.

Comparing Table 2 and Table 6 we observe differences between dropouts and non-dropouts. Below, in Table 3 there is a list of variables for which means and proportions in both panels are statistically different.

Table 3: Difference in variable means and proportions between Drop-outs and Non-dropouts (post-program surveys)

Variable name	Primary		Secondary	
	Female	Male	Female	Male
Enrollment rates	-0.450***	-0.443***	-0.508***	-0.531***
Percentage of children belonging to a poor family	3.5***	3.1***	3.7***	1.2
Mother's schooling (years)	-1.6***	-1.4***	-0.8***	-0.9***
Percentage of children with head of household employed	-2.6***	-2.1***	0.1	-1.9**
Percentage of children with father living at home	-3.5***	-6.1***	0.7	-0.2
Number of siblings enrolled at school	-0.7***	-0.6***	-0.6***	-0.6***
Community daily agricultural wage (pesos)	-1.2***	-0.4	1.7***	0.7**
Distance to secondary school (km)	0.9***	0.6***	0.9***	0.6***
Distance to metropolitan area (km)	-16.9***	-19.4***	-23.1***	-28.8***
Distance to the main city of her/his municipality (km)	-0.6**	-0.4**	-0.2	-0.8***

* Statistical significance = 10%. ** Statistical significance = 5%. *** Statistical significance = 1%.

The statistical significance of the differences is tested using tests for equality of means or proportions.

An important fact pointed out by Tables 2 and 3, is the low enrollment rate of dropouts. Only almost 60% of primary school children are actually attending class. Still worse is the situation for secondary children. Only 25% of them go to school. Compared with non-dropouts, enrollment rates after the implementation of the program are 45% lower for primary school dropouts, and more than 50% lower for secondary school dropouts. Some questions naturally arise from these figures. Why are these differences so large? Why is a child that decided not attend school once unlikely to re-enroll? Can we infer from these numbers that PROGRESA is not working all that well for dropouts contrary to what we expected?

To answer the first two questions take a look at Table 3. Again, as in Section 3.2, we conclude that, not considering grants, dropouts have more incentives to work than to attend school. Moreover, a higher proportion of dropouts come from poor families and with unemployed heads of households. Also, they have a higher direct cost of attending secondary school reflected in higher distances to secondary schools. Another explanation for the differences in enrollment rates could be that some unobserved characteristics as ability or personal ambition affect a child's schooling decision.

The remaining question is if the PROGRESA program is convincing those children who dropped out of school before the implementation of the program to go back to school and finish their education. If the answer is yes this implies that without the program enrollment rates would be much lower. On the other hand, if the program is not working for dropouts pre and post program enrollment rates should be equal. In both situations it is necessary to study alternatives schemes for the grant design that could send more dropouts back to school.

The answer cannot be obtained by just looking at descriptive statistics but needs to make use of the randomized assignment of the program. Comparison of results between treatment and control communities allows us to estimate the causal relationship between enrollment decision and PROGRESA grants.

5 Estimation of PROGRESA grants impact

5.1 Difference estimation

The random assignment of PROGRESA at community level has a crucial advantage. Randomization balances all observed and unobserved variables other than enrollment decision and treatment status across the two groups (treatment and control). Hence, this makes it possible to quantify the effect of the program on enrollment rates by simply comparing enrollment in treatment vs. enrollment in control communities, i.e., difference estimation can simply measure the program's effects.

To analyze if there exist differences in the effect of the program on non-dropout and dropouts we can estimate separately difference estimators for both groups, and compare the results.

In the context of this paper difference estimators are defined in the following form:

$$\hat{\mathbb{E}}_n(S_{it} \mid grant_{it} > 0, P_i = 1) - \hat{\mathbb{E}}_n(S_{it} \mid grant_{it} = 0, P_i = 1), \tag{1}$$

$$i = 1, ..., N \qquad t = 2, 3, 4$$

where S_{it}: dummy equal 1 if the child is enrolled in school at time t. $\hat{\mathbb{E}}_n$: post-program period averages. $grant_{it}$: the potential grant amount, that takes a value different from zero only if the child belongs to a poor family, resides in a PROGRESA community, and is attending a grade between 3^{rd} year of primary school and 3^{rd} year of junior secondary school. $grant_{it} > 0$ defines the treatment group while $grant_{it} = 0$ defines the control group[7]. P_i: is a dummy variable that takes on a value of 1 if the child belongs to a poor family. $t = 1, 2, 3, 4$ identify the October 1997 census and the October 1998, May 1999 and November 1999 surveys, respectively.

Unfortunately, difference estimation applied to the original data set partitioned in our two groups, dropouts and non-dropouts, may not be reliable. Randomization in the assignment of the program assures that any kind of analysis of the complete panel that implies disaggregation based on observable characteristics, other than the dependent variable and the treatment definition variable, is valid. Also, a sufficiently high number of observations is needed for a law of large numbers to hold in both groups defined by the treatment status. However, the variable that defines the groups under analysis is the dependent variable, school enrollment, in October 1997. Moreover, the dropouts panel fails to contain enough observations when we split the data between treatment and controls, by school level and by gender[8].

Only if randomization still holds when considering non-dropouts and dropouts observations separately, difference estimation is valid. But this is not the case here. I carried out an analysis of the randomization in both sub-panels following the methodology of Behrman and Todd (1999). Comparing means and distributions of observable characteristics between treatment and control observations I found some differences. Hence, the main conclusion is that the assignment of the program is not completely random when considering the groups as presented. Results for a set of relevant variables can be consulted in Table 9 and Table 10 of the Appendix. Therefore, difference estimation cannot provide accurate results. Nevertheless, I present the difference estimates of PROGRESA grants effect on Table 4 below and I compare these estimates with those obtained using a regression framework in Section 5.3.

[7]For children who fulfill the requirement to obtain the grant but are not receiving it I set $grant_{it} = 0$. In the calculus of difference estimates they belong to the control group.

[8]Size for each group are reported in Table 8 of the Appendix.

5.2 Regression framework

Following the discussion presented in Section 3.2 and including variables that reflect the impact of the PROGRESA program, a reduced form equation in latent variable form for the probability of being enrolled in school at time t, S^*_{it}, is[9]:

$$S^*_{it} = \eta_i + \alpha_{0t} + \alpha_1 P_i + \alpha_2 T_i + \sum_{k=2}^{8} \alpha_{3k} grant_{kit} + \sum_{k=1}^{K} \gamma_k C_{kit} + \sum_{j=1}^{J} \beta_j X_{jit} + e_{it} \quad (2)$$

$$i = 1, 2..., n \qquad t = 2, 3, 4 \qquad \text{and} \qquad e_{it} \sim F$$

What we observe, in fact is:

$$S_i = \mathbf{1}[S^*_i > 0], \qquad i = 1, 2..., n \quad (3)$$

η_i is an unobserved factor, individual specific and time-constant. It may reflect ability, personal ambition, etc. α_{0t} is a time variant unobserved effect. P_i is a dummy variable that takes on a value of 1 if the child belongs to a poor family. T_i is a dummy variable that takes on value 1 if the child lives in a community where the program started in September 1998, i.e., in a treatment community. $grant_{it}$ as defined in Section 5.1[10]. C_{kit} is equal to 1 if the child has successfully completed k years of school, $k = 1$ or less, $2, ..., 8$ and 9 or more, which qualifies the child for enrollment in $(k+1)^{th}$ grade. X_{jit} are a set of J individual, family, and community characteristics that includes the age of the child and the square of the age, mother's schooling, a dummy equal to one if the head of household was ill, a dummy equal to one if the head of household was employed in the week before the survey was conducted, a dummy set equal to 1 if the father lives at home, the number of girls younger than 16 years in the family, and the number of boys younger than 16 years in the family, the number of children younger than 5 years in the family, the number of adults women and men in the family, number of siblings enrolled in school, daily mean agricultural wage for men, distance to nearest junior secondary school, distance to nearest metropolitan area, and distance to the main city of her/his municipality. F is a distribution function.

[9]This reduced form equation is similar to the one proposed in Schultz (2004). The main differences are the introduction of an additional term to allow for time-constant unobserved effects and the introduction of a set of variables that allow for identification of differential effects of the program for dropouts.

[10]For children who fulfill the requirement to obtain the grant but are not receiving it I set $grant_{it} = 0$ and $T_i = 1$.

The expected values for the coefficients are the following. α_1 should be negative reflecting the common hypothesis that credit constraints limit the investment of the poor in their children's education. The effect of residing in a treatment community, or α_2, should be close to zero, since the assignment of the program is random, or slightly positive capturing some "spillover effects" of the treated communities on the control communities. α_{3k} captures the program effects, so it is greater than zero if the program successfully reaches its goal.

For the $\beta's$, we expect a negative effect of age, since for a given grade being older implies higher costs of schooling (higher opportunity costs for being more likely to get a job and to obtain a higher salary, psychological cost of disappointment if she/he failed, etc), a positive effect if the mother is more educated, a negative effect if the head of the household was ill and a positive effect if she/he had a job, a positive effect if the father lives at home, also a positive effect if the proportion of siblings attending school is higher, a negative effect from the opportunity cost of schooling (captured by wages), a negative effect from the direct cost of attending a junior secondary school (i.e. the non-existence of a school in the community), and finally a positive effect of the distance to the nearest metropolitan area and of the distance to the main city of her/his municipality.

To answer the question "Did PROGRESA send dropouts back to school?" it is necessary to model the probability of being enrolled for individual i at time t conditional on the schooling decision she/he made before the program started:

$$\mathbb{P}(S_{it} \mid S_{i1}), \qquad t = 2, 3, 4 \tag{4}$$

and then compare these probabilities between ex-ante dropouts ($S_{i1} = 0$) and children who were at school before the program started ($S_{i1} = 1$).

In order to capture the differences in the program's effects on non-dropouts compared to dropouts, the equation for the enrollment decision is modified as follows:

$$S^*_{it} = \eta_i + \alpha_{0t} + \alpha_1 P_i + \alpha_2 T_i + \alpha_3 D_i + \alpha_4 P_i * D_i + \alpha_5 T_i * D_i + \tag{5}$$
$$\sum_{k=2}^{8} \alpha_{6k} grant_{kit} + \sum_{k=2}^{8} \alpha_{7k} grant_{kit} * D_i + \sum_{k=1}^{K} \gamma_k C_{kit} + \sum_{j=1}^{J} \beta_j X_{jit} + e_{it},$$
$$i = 1, 2..., n \qquad \text{and} \qquad t = 2, 3, 4$$

where D_i is a dummy variable, that takes a value of 1 if the child dropped out of school before the program started. The impact of the program for non-dropouts is captured by the variable "grant", i.e., by the coefficient α_{6k}. The impact for dropouts is given by $\alpha_{6k} + \alpha_{7k}$. Hence, the difference in the program's impacts on non-dropouts and dropouts, is equal to α_{7k}.

In order to estimate the parameters of equation 5 we have to take into account its two main characteristics . First, it is a probability model, and second, there is an unobserved fixed effect.

A fixed effects conditional logit model would allow us to consistently estimate the parameters using a non-linear model and without any assumptions on how η_i is correlated with the exogenous variables. This approach is desirable for at least three reasons. The estimated probabilities are between 0 and 1, marginal effects are individual specific and it allows for the most flexible specification of the unobserved heterogeneity. However, such a model cannot be applied to the equation above since there is not enough variation in the data. A fixed effects non-linear estimation strategy can only considers observations for which the dependent variable has time variation. Applying this restriction to the PROGRESA panel we are left with 9,036 observations. Of those observations only 1,632 refer to dropouts. This is not enough data to identify the effect of interest.

Alternatively, we can use a fixed effects linear probability model (FELPM). It also has the most flexible specification for the fixed effect, so consistency is not an issue. Nevertheless, it has problems associated with the estimation of a probability using a linear model. One concern is about the marginal effects that in this model are assumed to be constant among individuals. In the context of schooling decision this assumption is not realistic. This feature may be overcome by allowing for non-linearities in the effect of the grant in the latent variable model, but only for time varying characteristics. Two important time-invariant factors affecting the schooling decision are the stock of education of the children's mother and the distance to the nearest school available, and by using a linear model the estimated marginal effects will be assumed to be the same for all children, regardless their mothers level of education and the transport cost they have to pay to attend school. We still can introduce interactions between children's characteristics relevant for the decision of attending school like age, schooling, and family composition variables, and most importantly by decomposing the effect of the grant by children's family characteristics. While this solution seems to be appealing for the estimation of the effect of the grant on non-dropout children, to decompose the effect of the grant on dropouts will rise the problem that some of these parameters will be identified by a small number of observations, so the accuracy of the resulting estimates may not be reliable. Another concern with the linear model is the number of observations for which the model estimates probabilities outside the interval $[0, 1]$. Using parameters values estimated with a FELPM, the proportion of observations for which the estimated probability lies outside the unit interval is 39%. This result makes difficult to asses how convincing is the fit of the model, that is usually tested comparing the distribution of actual and estimated probabilities.

Hence, I decided to estimate the effect of PROGRESA grants using the model proposed by Chamberlain (1980) andMundlak (1978), known as Correlated Random Effects Probit model (CREP). It is a non-linear probability model that impose some assumptions in the specification of the unobserved factor. In this sense, it is more restrictive than a fixed effects approach but also more flexible and its assumptions more likely to be fulfilled than a random effects model. The CREP model explicitly allows the individual specific unobservable term η_i to be correlated with time variant regressors assuming a conditional normal distribution with linear expectation and constant variance. The specification assumed for η_i is:

$$\eta_i = \psi + \xi \bar{x}_i + a_i, \tag{6}$$

where $\bar{x}_i$ is a vector including the average of: i)daily mean agricultural wage for men, ii)head of households' health and work status and iiii)grant amount interacted with the dropout dummy[11]

The complete set of assumptions for the identification of the parameters in the enrollment decision equation are the following:

1. $e_{it}|\eta_i, P_i, T_i, D_i, grant_{kit}, C_{kit}, X_{jit} \sim \Phi \qquad \forall i, \forall t, \forall k, \forall j$
   where Φ stands for the standard normal distribution

2. $S_{i1}...S_{iT}$ are independent conditional on $\eta_i, P_i, T_i, D_i, grant_{kit}, C_{kit}, X_{jit}$ $\forall i, \forall t, \forall k, \forall j$

3. $\eta_i|P_i, T_i, D_i, grant_{kit}, C_{kit}, X_{jit} \sim N(\psi + \bar{x}_i\xi, \sigma_a^2)$

To compute the average values of estimated marginal effects across treated individuals with the parameter's estimates obtained with the CREP model I proceed as

[11]This correlation is only allowed between time variant variables and the fixed effect. A potential source of bias in this model is the existence of correlation between the unobserved term and some time constant variables. In the model presented so far the dummy D_i may be correlated with η_i, since some unobserved factor could have determined the decision of dropping out of school. I considered this fact in the model including the grant amount multiplied by the dropout dummy in the vector $\bar{x}_i$.

follows.

Let Z_{kit} be the vector of explanatory variables:

$\overline{Z}_{kit} \equiv \left(1, \bar{x}_i, \overline{grant}_k, \overline{grant}_k * D_i, X_{1it}...X_{Jit}\right)'$

for each $k = 2...8$.

Let Z^0_{kit} be the same vector with the only difference that $grant_{ki} = 0$, for all individuals in all time periods:

$Z^0_{kit} \equiv (1, \bar{x}_i, 0, 0, X_{1it}...X_{Jit})'$

Let $\hat{\pi}_k$ be the vector of estimated parameters.

The average values of estimated marginal effects across treated individuals are calculated using the following expression:

$$\sum_{i:\ grant_i>0} \sum_{t=2}^{4} [\Phi(\hat{\pi}_k * \overline{Z}_{kit}) - \Phi(\hat{\pi}_k * Z^0_{kit})], \tag{7}$$

where Φ stands for the normal distribution function. I compute the average of the change in enrollment probabilities due to the implementation of the grant for children in conditions of receiving a grant (treated individuals). These averages are obtained for non-dropouts (i with $D_i = 0$) and dropouts (i with $D_i = 1$) separately.

5.3 Results

Table 4 presents the estimates of the effect of PROGRESA grants on school attendance for non-dropout and dropout children, by gender and education level, obtained with several estimation strategies. Notice that the results are reported grouping individuals according to the years of schooling completed as follows: from 2 to 5 years, corresponds to children in condition of attending primary school; a child that completed 6 years of schooling has graduated from primary school, and then is allowed to enter in secondary school; 7 and more years of schooling completed corresponds to children in condition of attending junior or secondary school. The last column of the table, column (11), shows the mean amount of the grant received by beneficiaries in the second semester of 1998 when they received the aid for the first time. These means are an approximation of the exact value used in the estimations, since the amount of the grant varies with the grade the child is attending and it was updated every semester to account for the increase in the level of general prices.

5.3.1 CREP model estimates: Main results

Columns (8) and (9) in Table 4 show average treatment effects on the treated computed using CREP model estimates of the parameters in the enrollment equation (equation 5). A complete report of these parameter's estimates can be found in Table 12 in the Appendix.

Consider non-dropout girls that have completed 6 years of schooling and are receiving the grant. The average probability of enrollment for a girl in this group is 5.3% higher when she is receiving the mean grant compared to when she is not receiving it. The probability of being enrolled is 5.3% higher due to the grant. With this kind of interpretation in mind, we can derive several conclusions.

In general the grant effect is positive. In four cases the effect is negative but insignificant due to huge standard errors, so the effect of interest in those cases is not clearly identified (as it happens with girls in primary school). The impact of the program is higher in secondary school than in primary school. This is an expected result because grants in secondary school are more than twice the amount of grants in primary school. Additionally, since enrollment rates are lower in the secondary level of education the program has more scope to work at this level.

The average effect of the grant for girls is higher than for boys in the non-dropout group. This is due to the fact that girls in secondary school are receiving higher grants than boys. Surprisingly, the same is not true in the group of dropouts. Drop-out boys react more strongly to the grant, even though they receive less money than girls.

The effects are different when we compare dropouts with non-dropouts. Since the standard errors of the estimated effects for dropouts in primary school are quite high, I do not made conclusions on these groups and, in what follows, all comments refers to secondary school children. For those children that have to enter in secondary school (6 years of schooling completed) the results are conclusive enough. There is no effect of grants in the re-enrolment decision of dropouts girls while for non-dropouts grants increase their enrolment probability by more than 5%. Drop-out boys react more to the incentive given by grants than non-dropouts. After receiving the grant the enrolment probability of both groups increase, but for dropouts this increase is almost 10% higher than for non-dropouts.

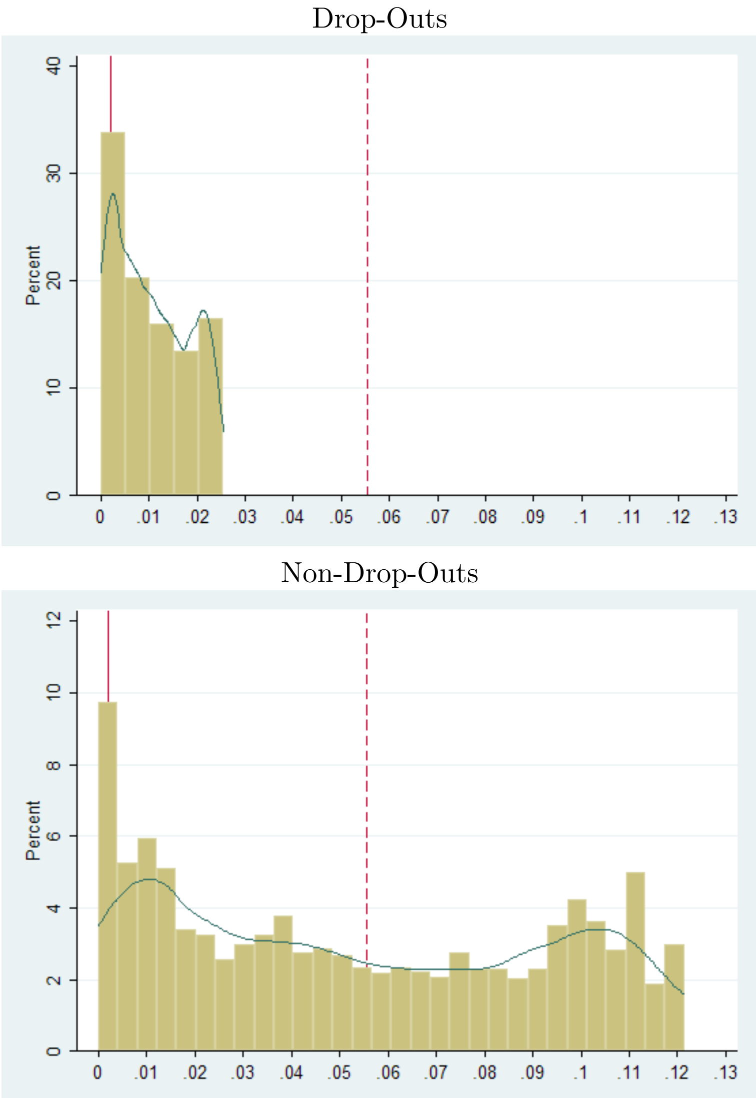
Table 4: Comparison of estimated marginal effects (percentage points)

Years of schooling completed			Difference Estimation			Regression Models with covariates							All Models
						Without unobserved heterogeneity Probit		With unobserved heterogeneity: Fixed effects Linear Probability Model		With unobserved heterogeneity: Correlated random effects Probit Model			Mean Grant in July 98' (mexican $)
			Marginal Effect (1)	Standard Error (2)	Sample Size (3)	Marginal Effect (4)	Standard Error (5)	Marginal Effect (6)	Standard Error (7)	Marginal Effect (8)	Standard Error (9)	Sample Size (10)	(11)
from 2 to 5	Female	Non-dropout	0.022***	(0.003)	14,435	0.009***	(0.003)	0.001	(0.010)	0.009***	(0.003)	8,118	108.8
		Dropout	0.022	(0.036)	724	-0.114***	(0.028)	-0.271**	(0.115)	-0.052	(0.115)	352	
		Difference	0.000	(0.015)		0.123***	(0.028)	0.272**	(0.115)	0.061	(0.115)		
	Male	Non-dropout	0.020***	(0.003)	15,725	0.002	(0.002)	0.015	(0.010)	0.002	(0.002)	8,893	108.8
		Dropout	0.152***	(0.032)	867	-0.021	(0.027)	0.046	(0.095)	0.107	(0.081)	474	
		Difference	-0.132***	(0.014)		0.023	(0.027)	-0.030	(0.096)	-0.106	(0.081)		
6	Female	Non-dropout	0.122***	(0.016)	3,567	0.055***	(0.014)	0.024	(0.020)	0.053***	(0.014)	1,889	209.2
		Dropout	0.126***	(0.020)	1,770	0.039	(0.025)	-0.286**	(0.140)	0.010	(0.112)	956	
		Difference	-0.004	(0.026)		0.015	(0.028)	0.310**	(0.141)	0.043	(0.112)		
	Male	Non-dropout	0.077***	(0.015)	3,760	0.019	(0.012)	0.015	(0.020)	0.019	(0.011)	2,050	199.2
		Dropout	0.046*	(0.024)	1,164	-0.039	(0.029)	0.055	(0.108)	0.114**	(0.058)	688	
		Difference	0.031	(0.030)		0.058*	(0.031)	-0.040	(0.109)	-0.095	(0.059)		
7 or more	Female	Non-dropout	0.024**	(0.008)	2,701	0.015**	(0.006)	0.011	(0.023)	0.016***	(0.006)	1,455	244.2
		Dropout	0.214***	(0.074)	188	-0.023	(0.048)	-0.331**	(0.145)	-0.196***	(0.068)	127	
		Difference	-0.190***	(0.036)		0.038	(0.047)	0.343**	(0.144)	0.212***	(0.067)		
	Male	Non-dropout	0.014	(0.009)	3,347	-0.002	(0.007)	0.010	(0.024)	-0.001	(0.007)	1,868	214.2
		Dropout	-0.017	(0.077)	166	0.011	(0.043)	0.026	(0.105)	-0.022	(0.115)	113	
		Difference	0.031	(0.043)		-0.013	(0.043)	-0.016	(0.106)	0.021	(0.115)		

Standard errors for difference estimates obtained by performing test of difference of proportions. Standard errors for regression models computed by bootstrap with 1000 replications. Cluster set at family level. Significance levels: $***$: 1% $**$: 5% $*$: 10%.
The set of covariates used in the regression models include child's information (age and education level), family characteristics (head of houselhold's health and working status, mother's education level, family composition and proportion of children attending school), and village characteristics (mean wage salary, distance to the main city at municipality level and to the nearest metropolitan area, and distance to the nearest village with a secondary school available). A detailed report of covariates can be seen in Table 11 in the Appendix.

Even though the equality of average marginal effects between dropouts and non-dropouts cannot be rejected at standard levels of significance the distribution of marginal effects between both groups show notable differences. Moreover, p-values from kolmogorov-smirnov tests of equality of distributions of marginal effects between dropouts and non-dropouts are always bellow 0.001, so the null hypothesis of equality is in all cases rejected even at 1% of significance. Figures 2 and 3 below depict marginal effects distributions stressing this conclusion.

Figure 2: Marginal effects distribution: female primary school completed (grade 6 completed)



- - - (dashed line) Average marginal effect for non-dropouts

—— (continuous line) Average marginal effect for dropouts

For girls that are about to enter in secondary school, the average impact is 5% for non-dropouts while the corresponding figure for dropouts is only 1% and statistically zero. Main characteristics of the distributions are a smaller range for dropouts, from 0.0 to 0.025, and for non-dropouts a high dispersion with almost all frequencies below 5%. Thus, for girls who have to decide whether to enter in secondary school grants are more convincing for those who were at school before the implementation of the program. PROGRESA education grants are not a strong enough incentive to persuade dropout girls to start secondary school. Notice that for this group the conclusion does not support the initial hypothesis that was made simply considering observable characteristics.

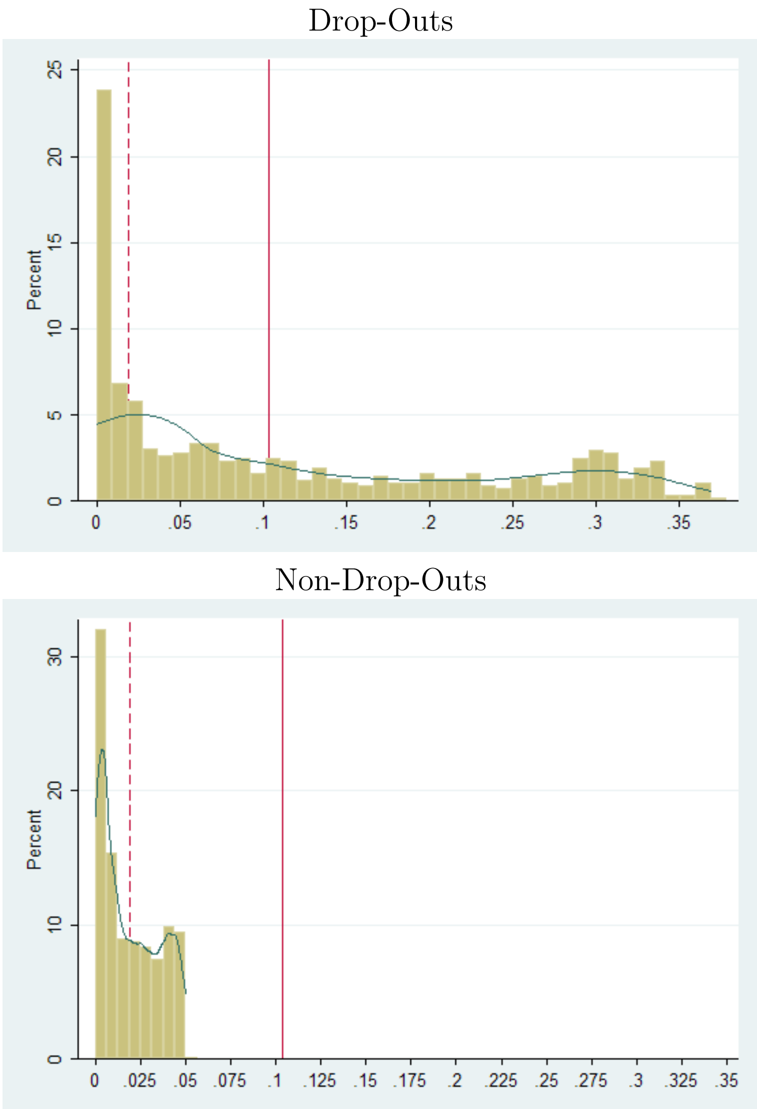
Turning to boys that have completed 6^{th} grade, there is an important and statistically significant difference in average marginal effects, more than 9% higher for dropouts (2% for non-dropouts, 11% for dropouts). For non-dropouts the distribution has a small range, from 0 to 0.05 with the highest frequency, more than 30%, between 0 and 0.002. The distribution for dropouts, on the other hand, is highly dispersed between 0 and 0.37 with frequencies in general below 5%. 25, 50 and 75 percentiles are all higher for dropouts. The message is clear. The program has a stronger effect on boys about to start secondary school who dropped-out in 1997 or before than on those who stayed on school.

Summing up, the estimated results presented above allow us to conclude that there exists a differential effect of the grant between children that dropped out of school before the program started and those who did not. The direction of differential effects in secondary school is not uniform for girls and boys. Drop-outs boys react more strongly to the program's incentives than non-dropouts. Non-dropouts girls have stronger grants' effects than dropouts. Therefore, I found the expected result of a higher effect of the grants on dropouts in secondary school only for boys. In general the program is effective for children that dropped out of school before they started receiving any grants. But when they have to decide to enter secondary school PROGRESA grants provide a better incentive for boys than for girls.

5.3.2 Estimates with alternative models

First, I report results of difference estimation. Columns (1), (2), and (3) in Table 4 show the average increase in school attendance due by PROGRESA grants, its estimated standard error and the number of observations used to compute the averages, respectively. The main difference with the results of the CREP model appears for children that have finished 6 years of education. The difference estimates show no differential effect of the grant between non-dropout and dropout girls, and a higher

Figure 3: Marginal effects distribution: male primary school completed (grade 6 completed)



- - - (dashed line) Average marginal effect for non-dropouts

—— (continuous line) Average marginal effect for dropouts effect of the grant for non-dropout boys than for dropout boys. As I discuss in Section 5.1 it is highly likely that these results are biased estimates of the true effect, since they do not control for differences in observable and unobservable characteristics between treated and control individuals.

Second, in columns (4) and (5) I present results of a probit model that controls for observed differences between treated and non-treated individuals, but still assumes that there are no differences in time invariant unobservables. With this model we would conclude that PROGRESA grants do not affect the probability of attending secondary school for dropout boys. However, it is difficult to accept estimates of the parameters in an equation reflecting school attendance without controlling for the existence of differences in ability. At a minimum, it is necessary to test if the estimated parameters do not change with the introduction of unobserved heterogeneity in the model.

Third, columns (6) and (7) report the average treatment effects on the treated and their corresponding standard errors estimated using a FELPM. The results are qualitatively consistent with those obtained with the CREP model. This provide evidence that the assumptions related with the unobserved heterogeneity in the CREP model are fulfilled, so the consistency of the estimates is validated. A report of the parameter estimates obtained with this model can be seen in Table in the Appendix.

5.3.3 Exploring the reasons behind the differential effect of PROGRESA grants

Ideally, one would want to estimate the effect of the grant for non-dropout and dropout children interacted with several observed characteristics to disentangle the reasons for the lack of effectiveness of PROGRESA grants in sending dropout girls back to secondary school. Unfortunately, the limited number of dropout observations in the sample forbids us to perform such exercise.

Still, I can exploit the nonlinear feature of the estimated marginal effects by computing the grant effect at specific values of observable characteristics. These results are reported in Table 13.

I will concentrate the analysis on the differential effect of the grant between non-dropouts and dropout children in condition of starting secondary school. Non-dropout children that have a numerous family (high number of siblings regardless of their gender and age) are more responsive the the grant, while this family characteristic reduces the effect of the grant for dropout children. In the case of girls, this finding may be rationalized in terms of their outside option. Only 12% of dropout girls who do not attend school report to be working. It is natural to thing that they may be staying at home helping in housework. Having a more numerous family makes them more necessary at home so they are less likely to go back to school.

The characteristics analyzed so far does not seem to help in explaining the differential effect of the grant between dropout girls and dropout boys.

Table 5: Average treatment effects evaluated at specific values of observable characteristics: children with 6 years of schooling completed

		Father living at home (yes minus no)		Mother's schooling (6 minus 2 years)		Distance to secondary school (3 km. minus 0 km.)	
		Effect	S.E.	Effect	S.E.	Effect	S.E.
Female	Non-dropout	-0.009***	(0.003)	-0.003**	(0.001)	0.003***	(0.001)
	Dropout	0.002	(0.021)	-0.001	(0.008)	-0.001	(0.006)
Male	Non-dropout	-0.005*	(0.003)	0.000	(0.000)	0.001	(0.001)
	Dropout	0.040*	(0.022)	-0.013	(0.009)	-0.010*	(0.006)

		Number of girls aged 6 or more (4 minus 0)		Number of boys aged 6 or more (4 minus 0)		Number of children aged 5 or less (4 minus 0)		Number of children attending school (5 minus 2)	
		Effect	S.E.	Effect	S.E.	Effect	S.E.	Effect	S.E.
Female	Non-dropout	0.032***	(0.008)	0.017***	(0.005)	0.005**	(0.002)	-0.044***	(0.011)
	Dropout	-0.007	(0.065)	-0.008	(0.079)	-0.001	(0.010)	0.000	(0.051)
Male	Non-dropout	0.007*	(0.004)	0.014*	(0.009)	0.002	(0.001)	-0.018*	(0.011)
	Dropout	-0.140*	(0.077)	-0.117	(0.076)	-0.016	(0.010)	0.055	(0.066)

Standard errors computed by bootstrap with 1000 replications. Cluster set at family level.
Significance levels: $***$: 1% $**$: 5% $*$: 10%.

6 Conclusions

I found evidence of the existence of a differential effect of PROGRESA grants on the school attendance decision of the overall target population and on the reentry decision of children who dropped out from school. These difference are observed in all groups analyzed. But the direction of the difference varies across groups. The expected result of a higher program effect on dropouts was found for boys in conditions of attending primary or secondary school. For girls in secondary school who dropped out in 1997 or before, the grant is not as good incentive to enroll in school as it is for non-dropout girls. Among dropouts in secondary school the impact of the education grants is lower for girls even though they receive more money than boys. The different responses of girls and boys to the grant in secondary school should be studied in more detail.

The last finding motivates the design of a particular model of schooling decision for girls. Individual variables such as marital state, pregnancy and number of children should be considered. Moreover, it can be argued that girls face a third option other than schooling or working. They may stay at home and take care of the children in the family. A model of schooling decisions should reflect this third option for girls. Additionally, there exist some perception that in the poorest rural communities girls are discriminated, so a monetary incentive like the PROGRESA grants may not be effective.

At a methodological level, the estimation of PROGRESA effects and the differential impact over dropouts can be improved by constructing a structural model of schooling. With the design of such a model, we might obtained a more conclusive answer to the question "Did PROGRESA send dropouts back to school?". Moreover, the estimation of a structural model will allow for the identification of a more effective and efficient policy that can send dropouts back to school.

References

Attanasio, O., C. Meghir, and A. Santiago (2005): "Education choices in Mexico: using a structural model and a randomized experiment to evaluate Progresa," *Institute for Fiscal Studies, Working Paper*, EWP05/01.

Behrman, J., and P. Todd (1999): "Randomness in the Experimental Samples of PROGRESA (Education, Health, and Nutrition Program)," *International Food Policy Research Institute, Washington, D.C. Research Report.*

Chamberlain, G. (1980): "Analysis of Covariance with Qualitative Data," *Review of Economic Studies*, 47, 225238.

Coady, D. P. (2000): "The application of social cost-benefit analysis to the evaluation of Progresa. Final Report," *International Food Policy Research Institute, Washington, D.C. Research Report.*

Mundlak, Y. (1978): "On the Pooling of Time Series and Cross Section Data," *Econometrica*, 46, 69–85.

Oportunidades (1999): "Reglas de operacion vigentes," *Secretaria de Desarrollo Social del Gobierno de Mexico downloadable from www.oportunidades.gob.mx.*

Schultz, P. (2004): "School subsidies for the poor: evaluating the Mexican Progresa poverty program," *Journal of Development Economics*, 74, 199–250.

Skoufias, E., B. Davis, and J. Behrman (1999a): "A. Final Report: An Evaluation of the Selection of Beneficiary Households in the Education, Health, and Nutrition Program (PROGRESA) of Mexico," *International Food Policy Research Institute, Washington, D.C. Research Report.*

——— (1999b): "B. An Addendum to the Final Report: An Evaluation of the Selection of Beneficiary Households in the Education, Health, and Nutrition Program (PROGRESA) of Mexico. Targeting the Poor in Mexico: Evaluation of the Selection of Beneficiary Households into PROGRESA," *International Food Policy Research Institute, Washington, D.C. Research Report.*

Todd, P., and K. I. Wolpin (2003): "Using a social experiment to validate a dynamic behavioral model of child schooling and fertility: assessing the impact of a school subsidy program in Mexico," *Penn Institute for Economic Research, Working Paper Series*, 03-022.

Appendix

Table 2*(continued)*

Variable	Primary		Secondary	
	Female	Male	Female	Male
Years of schooling completed				
0	0.000	0.003		
1	0.170	0.150		
2	0.114	0.117		
3	0.439	0.401		
4	0.113	0.154		
5	0.164	0.176		
6			0.863	0.816
7			0.071	0.072
8			0.019	0.039
9 or more			0.047	0.073
Age of child				
8	0.070	0.070		
9	0.115	0.090		
10	0.091	0.086		
11	0.092	0.089	0.001	0.000
12	0.106	0.097	0.015	0.015
13	0.125	0.097	0.096	0.087
14	0.127	0.133	0.253	0.209
15	0.134	0.149	0.310	0.306
16	0.111	0.144	0.247	0.288
17	0.029	0.045	0.075	0.093
18	0.002	0.001	0.004	0.002

Table 6: Descriptive statistics for Non-Drop-outs (post-program surveys)

Variable	Primary		Secondary	
	Female	Male	Female	Male
Sample size	25,564	23,521	8,659	9,735
Enrollment rate	0.976	0.971	0.767	0.790
Percentage of treatment communities	60.79	62.20	60.85	62.10
Percentage of children belonging to a poor family	89.44	89.85	80.51	81.87
Percentage of children eligible for receiving a grant	31.8	30.8	38.6	40.2
Grant (for grant different from zero) (pesos)	118.2	118.2	265.2	245.8
	(29.7)	(29.8)	(28.1)	(20.4)
Mother's schooling (years)	3.0	3.0	2.9	2.9
	(2.6)	(2.6)	(2.5)	(2.5)
Percentage of children with head of household ill	5.96	6.26	7.10	7.14
Percentage of children with head of household employed	91.80	91.60	90.03	89.51
Percentage of children with father not living at home	9.83	10.11	10.50	11.07
Number of girls from 5 to 16	2.0	1.0	1.9	0.9
	(1.2)	(1.0)	(1.1)	(1.0)
Number of boys from 5 to 16	1.0	1.9	1.0	1.9
	(1.0)	(1.1)	(1.0)	(1.1)
Number of children under 5	0.9	0.9	0.5	0.5
	(1.0)	(1.0)	(0.8)	(0.8)
Number of adult women	1.5	1.5	1.8	1.6
	(0.8)	(0.9)	(1.0)	(0.9)
Number of adult men	1.5	1.5	1.7	1.8
	(0.9)	(0.9)	(1.0)	(1.0)
Number of siblings enrolled at school	2.9	2.7	2.8	2.7
	(1.3)	(1.4)	(1.3)	(1.4)
Distance to secondary school (km)	2.3	2.3	1.9	2.0
	(2.1)	(2.1)	(1.8)	(1.8)
Percentage of children that have a secondary school in their community	23.2	26.5	32.6	28.3
Distance to nearest metropolitan area (km)	148.7	150.4	152.3	154.1
	(76.8)	(77.0)	(77.4)	(77.5)
Distance to the main city of her/his municipality (km)	11.7	11.6	11.8	11.5
	(8.1)	(7.9)	(8.2)	(8.0)
Community daily agricultural wage (pesos)	30.6	30.5	31.8	31.2
	(10.6)	(10.5)	(10.8)	(10.4)
Years of schooling completed				
0	0.000	0.000		
1	0.200	0.207		
2	0.181	0.184		
3	0.279	0.271		
4	0.177	0.172		
5	0.164	0.166		
6			0.489	0.455
7			0.231	0.245
8			0.175	0.190
9 or more			0.105	0.110
Age of child				
6	0.073	0.069		
7	0.146	0.136		
8	0.155	0.151		
9	0.155	0.152		
10	0.151	0.148		
11	0.151	0.142	0.023	0.021
12	0.092	0.098	0.181	0.139
13	0.042	0.056	0.255	0.224
14	0.022	0.028	0.232	0.247
15	0.008	0.013	0.182	0.210
16	0.003	0.005	0.103	0.127
17	0.000	0.001	0.022	0.030
18	0.000	0.000	0.002	0.001

Standard deviations are in parenthesis

Table 7: Comparison with Schultz (2004) confidence intervals for post-program difference estimation

Years of schooling completed		This paper 95% Conf. Interval		Schultz (2004)[a] 95% Conf. Interval	
2	Female	0.0144	0.0352	-0.0111	0.0471
	Male	0.0054	0.0217	0.0209	0.0210
3	Female	-0.0026	0.0155	-0.0122	0.0382
	Male	0.0061	0.0274	0.0489	0.0490
4	Female	0.0035	0.0267	0.0279	0.0481
	Male	0.0108	0.0332	0.0439	0.0440
5	Female	0.0287	0.0582	0.0457	0.0643
	Male	0.0101	0.0407	0.0409	0.0410
6	Female	0.0910	0.1521	0.1479	0.1480
	Male	0.0481	0.1054	0.0531	0.0769

[a]The results are taken from Table 3 in Schultz (2004).

Table 8: Groups size

Years of schooling completed in previous year	Non-dropouts				Drop-outs			
	Female		Male		Female		Male	
	Treatment	Control	Treatment	Control	Treatment	Control	Treatment	Control
From 2 to 5	8,118	6,317	8,893	6,832	352	372	474	393
6	1,889	1,678	2,050	1,710	956	814	688	476
7 or more	1,455	1,246	1,868	1,479	127	61	113	53

Table 9: p-values for test of randomization - Drop-out Observations

Variable	p-value (2,025 obs)	p-value based on community mean (440 obs)
Community population distribution	–	0.101 (3)
Distribution of communities over states	–	0.897 (3)
Age distribution of children under 16	0.892 (3)	0.925 (4)
Child's stock of education	0.107 (3)	0.340 (4)
Number of girls between 5 and 16 in the family	0.440 (3)	0.756 (4)
Number of boys between 5 and 16 in the family	0.426 (3)	0.336 (4)
Number of children under 5 in the family	0.099 (3)	0.856 (4)
Number of adult women in the family	0.002 (3)	0.456 (4)
Number of adult men in the family	0.180 (3)	0.720 (4)
Mother's schooling	0.000 (3)	0.955 (4)
Percentage of children with father not living at home	0.008 (2)	0.120 (4)
Percentage of children with head of household employed	0.019 (2)	0.321 (4)
Number of siblings enrolled at school	0.765 (1)	0.474 (4)
Community daily agricultural wage	0.000 (1)	0.914 (4)
Distance to metropolitan area	0.000 (1)	0.964 (4)
Distance to the main city of her/his municipality	0.001 (1)	0.742 (4)
Distance to secondary school	0.001 (1)	0.799 (4)

(1)Kolmogorov-Smirnov statistic for test of equality between two distribution functions. Ho: the distribution of the variable analyzed is equal in both groups
(2)T-test for equality of proportions. Ho: the variable analyzed has the same proportion of ones in both groups.
(3)Pearson's chi-squared statistic for the hypothesis that the frequencies in a two-way tabular are independent. Ho: the frequencies of the variable analyzed are independent.
(4)T-test for equality of means. Ho: the variable analyzed has the same mean in both groups.

Conclusion: At individual level several variables are different when comparing treatment and controls. At community level there exist statistical differences in a couple of variables. Hence, the random assignment of the program is lost when considering the group of dropouts separately.

Table 10: p-values for test of randomization - Non-dropout Observations

Variable	p-value (19,649 obs)	p-value based on community mean (492 obs)
Community population distribution	–	0.119 (3)
Distribution of communities over states	–	0.781 (3)
Age distribution of children under 16	0.230 (3)	0.925 (4)
Child's stock of education	0.009 (3)	0.369 (4)
Number of girls between 5 and 16 in the family	0.001 (3)	0.789 (4)
Number of boys between 5 and 16 in the family	0.000 (3)	0.415 (4)
Number of children under 5 in the family	0.000 (3)	0.295 (4)
Number of adult women in the family	0.000 (3)	0.701 (4)
Number of adult men in the family	0.007 (3)	0.525 (4)
Mother's schooling	0.002 (3)	0.930 (4)
Percentage of children with father not living at home	0.266 (2)	0.794 (4)
Percentage of children with head of household employed	0.016 (2)	0.324 (4)
Number of siblings enrolled at school	0.000 (1)	0.957 (4)
Community daily agricultural wage	0.000 (1)	0.997 (4)
Distance to metropolitan area	0.000 (1)	0.989 (4)
Distance to the main city of her/his municipality	0.000 (1)	0.414 (4)
Distance to secondary school	0.000 (1)	0.615 (4)

(1), (2), (3) and (4) idem Table 9.

Conclusion: There exist even more relevant differences at the individual level than those presented in Table 9. The evidence of lack of randomization is stronger when considering the group of non-dropouts separately.

Table 11: Description of variables in Tables 12 and 13

Variable name	Description
P	One if the child belongs to a poor family
T	One if the child resides in a community where PROGRESA grants were implemented in October 1998
D	One if the child was not enrolled in school in the October 1997 census
DP	Interaction between D (dropout) and P (poor family)
DT	Interaction between D (dropout) and T (treatment community)
grant	Grant amount (pesos)
grantd	Grant amount interacted with D (dropout child) (pesos)
wrepeon	Community daily agricultural wage for men (pesos)
health	One if the head of household was ill in the four weeks previous to the survey
work	One if the head of household has a job in the week previous to the survey
fhogar	One if the child's father is living at home with his family
distsec	Distance from the community where the child resides to the nearest community with a secondary school (km)
distmetro	Distance from the community where the child resides to the nearest metropolitan area. For communities in Hidalgo(state), these are Queretaro, Puebla, Tampico, or Mexico City; in Michoacan(state) it is Morelia; in Puebla it is Puebla; in Queretaro it is Queretaro; in San Luis de Potosi it is San Luis de Potosi; in Veracruz it is Veracruz and in Guerrero it is Acapulco (km)
distcab	Distance from the community where the child resides to the main city of her/his municipality (km)
schoolm	Years of schooling completed by the child's mother
girl	Number of girls from 5 to 16 years old in the child's family
boy	Number of boys from 5 to 16 years old in the child's family
baby	Number of children aged less than 5 years old in the child's family
women	Number of adult women (aged more than 16) in the child's family
men	Number of adult men (aged more than 16) in the child's family
w3	One for observations in the first post-program survey collected in October 1998
w4	One for observations in the second post-program survey collected in May 1999
age	Age of the child
age2	Square of the age of the child
ck	One if the child has completed k years of education (k = 1 or less, 2,..., 9 or more)
ageck	Age interacted with the stock of education of the child
asistest	Number of child's siblings enrolled in school
distsecd	Distance to secondary school interacted with dropout dummy D
distcabd	Distance to main city at municipality level interacted with dropout dummy D
workd	Variable work (head of household working status) interacted with dropout dummy D
schoolmd	Mother's stock of education interacted with dropout dummy D
workm	Time average for variable work (3 post-program observations)
healthm	Time average for variable health (3 post-program observations)
wrepeonm	Time average for variable wrepeon (3 post-program observations)
grantdm	Time average for variable grantd (3 post-program observations)
X_{fp}	Variable X interacted with a dummy variable equal 1 for girls in primary school
X_{fs}	Variable X interacted with a dummy variable equal 1 for girls in secondary school
X_{mp}	Variable X interacted with a dummy variable equal 1 for boys in primary school
X_{ms}	Variable X interacted with a dummy variable equal 1 for boys in secondary school
	X stands for the complete set of variables described above

Table 12: Probit estimates of enrollment probabilities[a]

		Number of obs =	74,427
		Wald $\chi^2(166)$ =	7,636.54
Log-pseudolikelihood =	-9565.1793	Prob > χ^2 =	0.0000
		Pseudo R^2 =	0.6688
Percentage correctly predicted = 95.34%			

enrolled	Coefficient	Standard Error	z	$P > \|z\|$	95% Conf.	Interval
P_{fp}	-0.06012	0.12335	-0.49	0.626	-0.30188	0.18164
T_{fp}	0.036389	0.096795	0.38	0.707	-0.15333	0.226104
D_{fp}	-1.78776	0.439293	-4.07	0.000	-2.64876	-0.92676
DP_{fp}	0.539533	0.345394	1.56	0.118	-0.13743	1.216493
DT_{fp}	0.233085	0.198013	1.18	0.239	-0.15501	0.621183
$grant_{fp}$	0.002198	0.000881	2.5	0.013	0.000471	0.003925
$grantd_{fp}$	-0.00452	0.003856	-1.17	0.242	-0.01207	0.003042
age_{fp}	0.329558	0.160167	2.06	0.04	0.015638	0.643479
$age2_{fp}$	-0.02601	0.007755	-3.35	0.001	-0.04121	-0.01081
$c2_{fp}$	2.228879	0.497638	4.48	0.000	1.253527	3.204232
$c3_{fp}$	1.676924	0.655976	2.56	0.011	0.391236	2.962613
$c4_{fp}$	1.73503	0.744903	2.33	0.02	0.275047	3.195013
$c5_{fp}$	1.829776	0.769984	2.38	0.017	0.320634	3.338917
$agec2_{fp}$	-0.22695	0.049548	-4.58	0.000	-0.32407	-0.12984
$agec3_{fp}$	-0.16305	0.057249	-2.85	0.004	-0.27525	-0.05084
$agec4_{fp}$	-0.14192	0.063374	-2.24	0.025	-0.26613	-0.01771
$agec5_{fp}$	-0.14879	0.06407	-2.32	0.02	-0.27436	-0.02322
$wrepeon_{fp}$	-0.00322	0.003184	-1.01	0.312	-0.00946	0.003019
$distsec_{fp}$	0.028863	0.016966	1.7	0.089	-0.00439	0.062116
$distcab_{fp}$	0.001253	0.00405	0.31	0.757	-0.00668	0.00919
$distmetro_{fp}$	0.001298	0.000486	2.67	0.008	0.000346	0.00225
$schoolm_{fp}$	-0.00521	0.014298	-0.36	0.715	-0.03324	0.02281
$health_{fp}$	-0.21307	0.083881	-2.54	0.011	-0.37747	-0.04867
$work_{fp}$	-0.09014	0.113698	-0.79	0.428	-0.31298	0.132706
$fhogar_{fp}$	0.40628	0.137486	2.96	0.003	0.136812	0.675747
$girl_{fp}$	-0.44548	0.048201	-9.24	0.000	-0.53995	-0.35101
boy_{fp}	-0.53084	0.04877	-10.88	0.000	-0.62642	-0.43525
$baby_{fp}$	-0.00321	0.034287	-0.09	0.925	-0.07041	0.063991
$women_{fp}$	0.005986	0.047786	0.13	0.9	-0.08767	0.099645
men_{fp}	0.024404	0.038311	0.64	0.524	-0.05068	0.099491
$asistest_{fp}$	1.505614	0.059769	25.19	0.000	1.38847	1.622758
$distsecd_{fp}$	-0.05231	0.03061	-1.71	0.087	-0.1123	0.007689
$distcabd_{fp}$	0.019577	0.009987	1.96	0.05	2.31E-06	0.039151
$workd_{fp}$	0.535823	0.254472	2.11	0.035	0.037067	1.034579
$grantdm_{fp}$	-0.00262	0.004232	-0.62	0.536	-0.01091	0.005676
$wrepeonm_{fp}$	0.006664	0.004662	1.43	0.153	-0.00247	0.015802
$workm_{fp}$	0.28682	0.195492	1.47	0.142	-0.09634	0.669977
$healthm_{fp}$	0.521563	0.215596	2.42	0.016	0.099003	0.944122
$workdm_{fp}$	-0.35971	0.353035	-1.02	0.308	-1.05165	0.332224
$asistemst_{fp}$	-0.66512	0.053696	-12.39	0.000	-0.77037	-0.55988
$w3_{fp}$	0.093972	0.049281	1.91	0.057	-0.00262	0.19056
$w4_{fp}$	-0.05902	0.044302	-1.33	0.183	-0.14585	0.027812
$cons$	0.17941	0.851053	0.21	0.833	-1.48862	1.847444

(continued on next page)

[a] Variable's definition are explained in Table 11.

Table 12 *(continued)*

enrolled	Coefficient	Standard Error	z	$P > \|z\|$	95% Conf.	Interval
P_{fs}	-0.33882	0.095044	-3.56	0.000	-0.5251	-0.15254
T_{fs}	0.120591	0.088996	1.36	0.175	-0.05384	0.29502
D_{fs}	8.186298	5.750446	1.42	0.155	-3.08437	19.45696
DP_{fs}	0.075007	0.202451	0.37	0.711	-0.32179	0.471803
DT_{fs}	-0.0103	0.197861	-0.05	0.958	-0.3981	0.377501
$grantc6_{fs}$	0.001144	0.000435	2.63	0.009	0.000292	0.001997
$grantdc6_{fs}$	-0.0009	0.002183	-0.41	0.679	-0.00518	0.003375
$grantc7_{fs}$	0.001022	0.000473	2.16	0.031	9.41E-05	0.00195
$grantdc7_{fs}$	-0.00545	0.001853	-2.94	0.003	-0.00908	-0.00182
age_{fs}	-0.71461	0.461384	-1.55	0.121	-1.61891	0.189682
$age2_{fs}$	0.016453	0.014813	1.11	0.267	-0.01258	0.045486
$c6_{fs}$	1.629472	1.315147	1.24	0.215	-0.94817	4.207113
$c7_{fs}$	5.006646	1.351812	3.7	0.000	2.357142	7.656149
$agec6_{fs}$	-0.11693	0.086412	-1.35	0.176	-0.28629	0.052434
$agec7_{fs}$	-0.24183	0.088076	-2.75	0.006	-0.41445	-0.0692
$wrepeon_{fs}$	0.004344	0.00267	1.63	0.104	-0.00089	0.009577
$distsec_{fs}$	-0.05119	0.013887	-3.69	0.000	-0.07841	-0.02397
$distmetro_{fs}$	0.002178	0.000368	5.91	0.000	0.001456	0.0029
$distcab_{fs}$	0.000799	0.003078	0.26	0.795	-0.00523	0.006831
$schoolm_{fs}$	0.032229	0.013253	2.43	0.015	0.006254	0.058204
$health_{fs}$	-0.15281	0.065867	-2.32	0.02	-0.28191	-0.02372
$work_{fs}$	0.008876	0.079652	0.11	0.911	-0.14724	0.164991
$fhogar_{fs}$	0.512864	0.123557	4.15	0.000	0.270697	0.755031
$girl_{fs}$	-0.64605	0.05665	-11.4	0.000	-0.75708	-0.53501
boy_{fs}	-0.69022	0.057308	-12.04	0.000	-0.80254	-0.5779
$baby_{fs}$	-0.06231	0.028582	-2.18	0.029	-0.11833	-0.0063
$women_{fs}$	-0.00992	0.027727	-0.36	0.72	-0.06427	0.04442
men_{fs}	-0.0307	0.024896	-1.23	0.217	-0.0795	0.018092
$asistest_{fs}$	1.651837	0.065212	25.33	0.000	1.524023	1.779651
$aged_{fs}$	-1.2938	0.793451	-1.63	0.103	-2.84893	0.261335
$age2d_{fs}$	0.046772	0.027289	1.71	0.087	-0.00671	0.100257
$schoolmd_{fs}$	-0.07881	0.031583	-2.5	0.013	-0.14071	-0.01691
$grantdc6m_{fs}$	0.000432	0.002275	0.19	0.849	-0.00403	0.004891
$grantdc7m_{fs}$	0.008237	0.002846	2.89	0.004	0.002658	0.013815
$wrepeonm_{fs}$	-0.00456	0.003939	-1.16	0.247	-0.01228	0.00316
$workm_{fs}$	0.286502	0.175365	1.63	0.102	-0.05721	0.630211
$healthm_{fs}$	0.244266	0.215016	1.14	0.256	-0.17716	0.66569
$asistemst_{fs}$	-0.67138	0.047703	-14.07	0.000	-0.76488	-0.57789
$w3_{fs}$	0.378933	0.049531	7.65	0.000	0.281854	0.476012
$w4_{fs}$	0.35171	0.042589	8.26	0.000	0.268237	0.435182
fs	5.345003	3.769465	1.42	0.156	-2.04301	12.73302

(continued on next page)

Table 12 *(continued)*

enrolled	Coefficient	Standard Error	z	$P > \|z\|$	95% Conf.	Interval
P_{mp}	0.0019842	0.1115879	0.02	0.986	-0.216724	0.2206924
T_{mp}	0.1787204	0.0759595	2.35	0.019	0.029843	0.3275982
D_{mp}	3.9781890	2.4427500	1.63	0.103	-0.809512	8.7658900
DP_{mp}	-0.2495031	0.2543473	-0.98	0.327	-0.748015	0.2490085
DT_{mp}	0.2368955	0.1807437	1.31	0.190	-0.117356	0.5911466
$grant_{mp}$	0.0003683	0.0007122	0.52	0.605	-0.001028	0.0017642
$grantd_{mp}$	0.0045421	0.0035240	1.29	0.197	-0.002365	0.0114490
age_{mp}	0.4517245	0.1438856	3.14	0.002	0.169714	0.7337351
$age2_{mp}$	-0.0293674	0.0068852	-4.27	0.000	-0.042862	-0.0158727
$c2_{mp}$	2.0547600	0.4078813	5.04	0.000	1.255328	2.8541930
$c3_{mp}$	2.7626040	0.5830192	4.74	0.000	1.619907	3.9053000
$c4_{mp}$	2.6819920	0.6220860	4.31	0.000	1.462726	3.9012580
$c5_{mp}$	2.1837420	0.7014966	3.11	0.002	0.808834	3.5586500
$agec2_{mp}$	-0.1899496	0.0392167	-4.84	0.000	-0.266813	-0.1130863
$agec3_{mp}$	-0.2544400	0.0485934	-5.24	0.000	-0.349681	-0.1591988
$agec4_{mp}$	-0.2101065	0.0499493	-4.21	0.000	-0.308005	-0.1122077
$agec5_{mp}$	-0.1773750	0.0550845	-3.22	0.001	-0.285339	-0.0694115
$wrepeon_{mp}$	-0.0038323	0.0033377	-1.15	0.251	-0.010374	0.0027095
$distsec_{mp}$	-0.0003181	0.0134879	-0.02	0.981	-0.026754	0.0261176
$distmetro_{mp}$	0.0025447	0.0004920	5.17	0.000	0.001580	0.0035091
$distcab_{mp}$	0.0006364	0.0038155	0.17	0.868	-0.006842	0.0081146
$schoolm_{mp}$	0.0178968	0.0132084	1.35	0.175	-0.007991	0.0437848
$health_{mp}$	0.0215211	0.1050984	0.20	0.838	-0.184468	0.2275102
$work_{mp}$	0.0301827	0.0868495	0.35	0.728	-0.140039	0.2004046
$fhogar_{mp}$	0.4758867	0.1065015	4.47	0.000	0.267148	0.6846258
$girl_{mp}$	-0.5780486	0.0397477	-14.54	0.000	-0.655953	-0.5001446
boy_{mp}	-0.4682509	0.0404441	-11.58	0.000	-0.547520	-0.3889820
$baby_{mp}$	0.0172073	0.0294409	0.58	0.559	-0.040496	0.0749103
$women_{mp}$	0.0573411	0.0347276	1.65	0.099	-0.010724	0.1254059
men_{mp}	0.0066312	0.0296029	0.22	0.823	-0.051390	0.0646518
$asistest_{mp}$	1.6161750	0.0665607	24.28	0.000	1.485718	1.7466310
$aged_{mp}$	-0.7478791	0.3750563	-1.99	0.046	-1.482976	-0.0127824
$age2d_{mp}$	0.0271769	0.0145600	1.87	0.062	-0.001360	0.0557140
$wrepeond_{mp}$	0.0129225	0.0072253	1.79	0.074	-0.001239	0.0270837
$schoolmd_{mp}$	-0.1354224	0.0340484	-3.98	0.000	-0.202156	-0.0686886
$grantdm_{mp}$	-0.0056409	0.0036899	-1.53	0.126	-0.012873	0.0015912
$wrepeonm_{mp}$	0.0077109	0.0044644	1.73	0.084	-0.001039	0.0164610
$workm_{mp}$	-0.2277633	0.1669524	-1.36	0.172	-0.554984	0.0994573
$healthm_{mp}$	0.1251161	0.2103195	0.59	0.552	-0.287103	0.5373348
$wrepeondm_{mp}$	-0.0077973	0.0093024	-0.84	0.402	-0.026030	0.0104351
$asistem_{s}t_{mp}$	-0.6956268	0.0626527	-11.10	0.000	-0.818424	-0.5728298
$w3_{mp}$	0.1979129	0.0510996	3.87	0.000	0.097760	0.2980662
$w4_{mp}$	0.0023492	0.0438664	0.05	0.957	-0.083627	0.0883257
mp	-1.0594210	1.1248610	-0.94	0.346	-3.264109	1.1452670

(continued on next page)

Table 12 *(continued)*

enrolled	Coefficient	Standard Error	z	$P > \|z\|$	95% Conf.	Interval
P_{ms}	-0.35897	0.079264	-4.53	0.000	-0.51433	-0.20362
T_{ms}	0.119921	0.077159	1.55	0.12	-0.03131	0.27115
D_{ms}	-1.11906	0.215512	-5.19	0.000	-1.54146	-0.69667
DP_{ms}	0.354124	0.207761	1.7	0.088	-0.05308	0.761328
DT_{ms}	0.263892	0.204407	1.29	0.197	-0.13674	0.664523
$grantc6_{ms}$	0.000491	0.000407	1.21	0.227	-0.00031	0.001289
$grantdc6_{ms}$	0.003254	0.00192	1.69	0.09	-0.00051	0.007018
$grantc7_{ms}$	-5.1E-05	0.000481	-0.11	0.916	-0.00099	0.000893
$grantdc7_{ms}$	-0.00036	0.001714	-0.21	0.833	-0.00372	0.002997
age_{ms}	-0.80319	0.406263	-1.98	0.048	-1.59945	-0.00693
$age2_{ms}$	0.019289	0.012876	1.5	0.134	-0.00595	0.044525
$c6_{ms}$	2.651963	1.256475	2.11	0.035	0.189318	5.114608
$c7_{ms}$	3.685036	1.233131	2.99	0.003	1.268144	6.101929
$agec6_{ms}$	-0.17643	0.081999	-2.15	0.031	-0.33714	-0.01571
$agec7_{ms}$	-0.16052	0.079794	-2.01	0.044	-0.31691	-0.00412
$wrepeon_{ms}$	-0.00493	0.002574	-1.91	0.056	-0.00997	0.00012
$distsec_{ms}$	-0.04974	0.01343	-3.7	0.000	-0.07606	-0.02341
$distmetro_{ms}$	0.002609	0.000382	6.82	0.000	0.001859	0.003358
$distcab_{ms}$	0.000743	0.003112	0.24	0.811	-0.00536	0.006842
$schoolm_{ms}$	0.014608	0.011231	1.3	0.193	-0.0074	0.036619
$health_{ms}$	0.021708	0.08097	0.27	0.789	-0.13699	0.180406
$work_{ms}$	0.106334	0.07132	1.49	0.136	-0.03345	0.246118
$fhogar_{ms}$	0.682057	0.11782	5.79	0.000	0.451135	0.912979
$girl_{ms}$	-0.8378	0.052616	-15.92	0.000	-0.94092	-0.73467
boy_{ms}	-0.72726	0.050609	-14.37	0.000	-0.82645	-0.62807
$baby_{ms}$	-0.06182	0.027068	-2.28	0.022	-0.11487	-0.00876
$women_{ms}$	0.000452	0.029253	0.02	0.988	-0.05688	0.057786
men_{ms}	-0.05032	0.026325	-1.91	0.056	-0.10191	0.001278
$asistest_{ms}$	1.83169	0.062904	29.12	0.000	1.708399	1.95498
$schoolmd_{ms}$	-0.06732	0.032089	-2.1	0.036	-0.13021	-0.00442
$grantdc6m_{ms}$	-0.00508	0.00197	-2.58	0.01	-0.00894	-0.00121
$grantdc7m_{ms}$	0.002296	0.002065	1.11	0.266	-0.00175	0.006343
$wrepeonm_{ms}$	0.000636	0.003678	0.17	0.863	-0.00657	0.007845
$workm_{ms}$	0.056352	0.155714	0.36	0.717	-0.24884	0.361547
$healthm_{ms}$	-0.28532	0.187619	-1.52	0.128	-0.65305	0.082409
$asistemst_{ms}$	-0.73616	0.052679	-13.97	0.000	-0.83941	-0.63291
$w3_{ms}$	0.267694	0.047753	5.61	0.000	0.174099	0.361288
$w4_{ms}$	0.223197	0.040794	5.47	0.000	0.143243	0.303151
ms	6.435053	3.395908	1.89	0.058	-0.22081	13.09091

Table 13: Fixed-effects (within) regression coefficients[a]

				Number of obs =		74,427
				Number of groups =		24,809
R-sq:	within =	0.3887		Obs per group:		min = 3
	between =	0.3643				avg = 3
	overall =	0.3580				max = 3
$F(86, 49{,}532) =$	366.15					
corr(u_i,Xb) =	-0.4034				Prob $>F=$	0.0000
enrolled	Coefficient	Standard Error	t	$P>\|t\|$	95% Conf.	Interval
$grant_{fp}$	0.0000	0.0001	0.0900	0.9250	-0.0001	0.0002
$grantd_{fp}$	-0.0023	0.0005	-4.9700	0.0000	-0.0032	-0.0014
$grantc6_{fs}$	0.0001	0.0001	1.6900	0.0900	0.0000	0.0002
$grantd6_{fs}$	-0.0013	0.0002	-5.4300	0.0000	-0.0017	-0.0008
$grantc7_{fs}$	0.00003	0.0001	0.7600	0.4480	-0.0001	0.0001
$grantd7_{fs}$	-0.0012	0.0002	-6.0500	0.0000	-0.0016	-0.0008
$grant_{mp}$	0.0001	0.0001	1.6600	0.0970	0.0000	0.0003
$grantd_{mp}$	0.0002	0.0004	0.5700	0.5700	-0.0006	0.0011
$grantc6_{ms}$	0.0001	0.0001	1.1200	0.2640	0.0000	0.0002
$grantd6_{ms}$	0.0002	0.0002	0.7300	0.4660	-0.0003	0.0006
$grantc7_{ms}$	0.00003	0.0001	0.7300	0.4660	-0.0001	0.0001
$grantd7_{ms}$	0.0001	0.0002	0.2900	0.7700	-0.0004	0.0005
σ_u	0.2598					
σ_e	0.1574					
ρ	0.7315	(fraction of variance due to u_i)				
F test that all $u_i=0$: $F(24{,}808,\ 49{,}532) = 4.37$ Prob $>F=0.0000$						

Standard errors are clustered at family level. [a] Variable's definition are explained in Table 11.