The school reentry decision of poor girls. Structural estimation and policy analysis using PROGRESA database.

María Nieves Valdés*†
Universidad de Santiago de Chile

PRELIMINARY VERSION

Abstract

In this paper I present a dynamic structural model of girls' schooling choices and estimate it using the Mexican PROGRESA database. This structural approach allows evaluating the effectiveness of several policies to increase school reentry rates for girls in low-income households. To increase school attendance among poor children in developing countries, policy makers have implemented conditional cash transfers programs. Although transfers have been highly successful in keeping girls at school, they have a lower and relatively small effect in increasing school attendance among girls who have dropped out of school. Cash transfer programs may fail because most of these poor girls leave school to stay at home helping in housework, rather than working for a salary. Results suggest that effective policies to increase school reentry rates for poor girls are sizable increases in the amount of the grant, free access to community nurseries and kindergartens, and generalizing the availability of secondary schools.

JEL classification I21; I28; J16; O15

Keywords Policy Evaluation; Dynamic discrete choice structural models; School choices for girls; School reentry; PROGRESA

*Email: nieves.valdes@usach.cl.

†Special thanks to Pedro Albarrán and Ricardo Mora for their advice, remarks and comments. I am grateful to Julio Cáceres, Jesús Carro, Irma Clots, Ramiro de Elejalde, Dolores de la Mata, Sara de la Rica, Pedro Mira, 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, the 11th IZA European Summer School in Labor Economics, ESPE 2008 Meeting, XXXIII Simposio de la Asociación Española de Economía, Central European University, Universidad de Islas Baleares, Meeting of the Network of Inequality and Poverty, Universidad de Chile, Department of Economics, Universidad de Chile, EALE/SOLE 2010 3^{rd} International Conference, Department of Economics, Universidad de Santiago de Chile, and 3rd IZA Workshop on Child Labor in Developing Countries.

1 Introduction

In this paper I evaluate the effectiveness of alternative policies to persuade girls who dropped out of school and belong to poor families to return to school and continue with their education. I discuss the differential effect that several policies have on reentry decisions and on enrollment decisions. I quantify the effect of demand-side policies such as conditional cash transfers and availability of daycare centers for young children. I also present results of the effect of a supply-side policy on school attendance that is the increase in the number of communities where a secondary school is available. The analysis is based on a dynamic behavioral model of school choices for girls. The structural parameters of the model are estimated using the Mexican PROGRESA database.

The motivation behind this study is threefold. First, in the paper I address a relevant policy concern: how to increase educational participation in developing countries. Despite the efforts made by policy makers in increasing enrollment rates, educational participation is far from targets proposed by several international institutionsreports that between 2000 and 2006, the total number of out-of-school children in low-income countries decreased by 41%.[1]. UNESCO (2007) However, in 2006 almost one in five of children of primary school age was not in school. Secondary net enrolment rates have been gradually increasing by around 2 to 3 percentage points per year in most regions. Still, in 2005, three in five children of secondary school age in low-income countries were not in school. To increase school attendance rates among poor children in developing countries, policy makers have implemented conditional cash transfers programs.[2] Although transfers have been successful in keeping boys and girls at school, there exist evidence that they do not increase girls' reentry rates.

Second, in this paper I contribute with the evaluation of a well-known anti-poverty program, PROGRESA. I quantify the effect of PROGRESA grants on the girls' reentry decision using a dynamic behavioral model of school choices. The methodology applied allows performing a counterfactual analysis. The analysis of the effect of PROGRESA grants on reentry decision

[1]For example, universal primary education is Goal 2 of both Education for All movement and the Millennium Development Goals adopted by UN Member States in 2000

[2]Examples of cash transfer programs are PROGRESA in Mexico, PRAF in Honduras, Red de Protección Social in Nicaragua, and Familias en Acción in Colombia.

has been seldom discussed. Behrman, Sengupta, and Todd (2001), using difference-in-difference estimation techniques, conclude that PROGRESA grants increase reentry rates and this effect is lower for girls than for boys. Valdes (2008) addresses the analysis of the effect of PROGRESA grants on reentry rates by estimating a reduced form equation for schooling enrollment and finds that grants increase reentry rates among boys but do not affect girls' reentry rates.

Third, this study contributes to a growing literature that addresses empirical questions using discrete choice dynamic programming models of individual behavior. These models are attractive because structural parameters have a clear interpretation within the theoretical model and they are useful tools for the evaluation of counterfactual policies (Aguirregabiria and Mira (2007)). Miller (1984) and Keane and Wolpin (1997) propose and estimate dynamic models of occupational choices. Attanasio, Meghir, and Santiago (2005) and Todd and Wolpin (2006) use dynamic behavioral models to evaluate the PROGRESA program.

In this paper schooling choices for girls in poor families are modeled following the individual decision approach used by Attanasio, Meghir, and Santiago (2005), where boys decide whether to attend school or to work. For families with many children the value of retaining a girl at home becomes more relevant because they are a good help in housework. As girls may dropout from school to stay at home I depart from Attanasio, Meghir, and Santiago (2005) by allowing girls to choose among three alternatives: attend school, stay at home, and work. Under this framework a girl schooling decision can be assumed to be made by her parents in an altruistic fashion. That is, they choose the alternative that maximizes their daughter's inter-temporal welfare independently of the decision they make for their other children. I relax the assumption of allowing the value of each alternative to be affected by family composition in two ways. Unobserved individual heterogeneity and the utility a girl derives from staying at home are affected by family characteristics.

The estimated model fits girls' schooling choices reasonably well. It replicates patterns observed in the actual distribution of schooling choices by ages: for each particular age, reentry rates are lower than enrollment rates; reentry and enrollment rates decrease as age increases, and reentry rates decrease faster than enrollment rates. It also replicates main features of the distribution of schooling choices by stock of education: reentry and enrollment rates decreases as the stock of education increases and, in the last grade of primary school and in the last grade of junior secondary school reentry and enrollment rates go down remarkably. It is observed in the data that most girls who were attending school in the previous academic year (non-dropouts) were still in school in the next academic year, whereas only 40% of girls who were out of school (dropouts) came back. The estimated model is able to match these differences in the distribution of schooling choices between female non-dropouts and dropouts. It rationalizes these differences by showing that persistence and unobserved ability at school are relevant in the decision to attend school. The model also contributes to understand the reasons that make a girl dropout from school. A girl's decision to drop out of school is related to her age, the composition of her family, her mother's labor participation, and also related to unobserved characteristics of the girl, such as her unobserved ability at school. As her value at home increases with the number of members in her family and with her age, she leaves school not to work but to stay at home helping in housework. Additionally, results suggest that alternative policies to cash transfers, such as free access to community nurseries and kindergartens, and increasing the availability of secondary schools, effectively increase school reentry rates for poor girls.

The paper is organized as follows. Section 2 presents the main features of the PROGRESA program and describes the characteristics of the PROGRESA database. It also provides some main statistics that focus on the differences between dropouts and non-dropouts. In Section 3, I present the theoretical model and discuss its empirical implementation. In Section 4, I present the results of the estimation of the structural parameters, I discuss the validity of the estimated parameters, and I show the results of the counterfactual analysis. Finally, Section 5 concludes the paper with its main results.

2 The Data

2.1 Description of PROGRESA

The Education, Health and Nutrition program, PROGRESA, was first 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 aims 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 expected outcomes were higher literacy rates, enrollment rates, and completion rates; lower child mortality rates and higher vaccination rates; and lower rates of undernourishment. The program is targeted at the family level. 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 characteristics such as presence of running water, electricity, pipes, etc.[3] Eligibility is independent of residence and family size and composition. All aid is given to the mother as there is evidence that mothers are better than fathers at allocating family resources.[4]

The education component includes monthly grants for children of a family qualified as beneficiary. To be given a grant, children need to be less than 18 years old, enrolled in school between the third year of primary school and the third year of junior secondary school, and to fulfill a minimum attendance requirement. The grants are not assigned based on academic achievement. A child who does not pass a grade is still eligible for the grant in the following year. However, if the child fails the same grade twice, she/he losses eligibility. The grant increases with the years of schooling completed. In the junior secondary level the grant is slightly higher for girls, because there is evidence that in poor families girls are more likely to dropout of school and that they dropout earlier than boys. Additionally, beneficiaries receive an annual grant for school supplies. In Table 1, there is a description of grants amounts. An eligible family was entitled to receive at most 420 pesos per month by means of scholarships in the second half of 1998. This amount represents 40% of the mean monthly family income and 67% of the mean monthly family expenditure in consumption. Thus, scholarships are potentially an important source of household resources.

2.2 Evaluation 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.

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

[4]See Rubalcava and Thomas (2000) for a discussion.

Table 1: Grant amount and household income and consumption (in Mexican pesos)

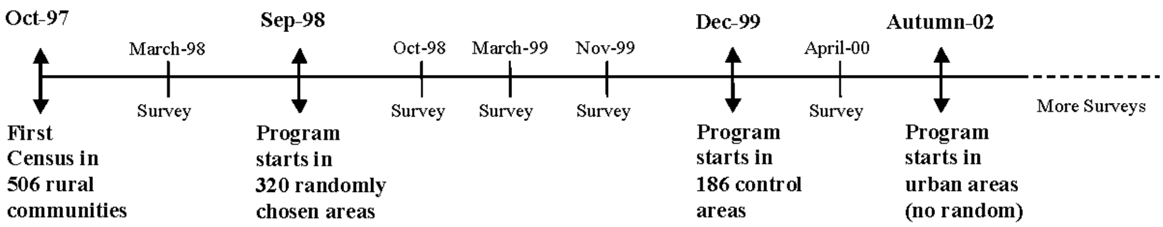
Monthly grant		July - Dec, 98	Jan - June,99
Primary School			
	3	70	75
	4	80	90
	5	105	115
	6	135	150
Secondary School			
1	girl	195	210
	boy	185	200
2	girl	220	235
	boy	195	210
3	girl	240	255
	boy	205	225
Monthly maximum support by means of grants per family		420	465
Annual aid for school supplies			Academic year 98/99
Primary School			135
Secondary School			170

Monthly Household Income and Consumption	Nov 98
Income	1071
Consumption	630

Source: Data on grants from Histórico de apoyos monetarios. SEDESOL 2005. Data on income and consumption from Albarran and Attanasio (2002)

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 afterward. In October 1998, the program was implemented in 320 randomly selected communities (treated communities), whereas in the remaining 186 communities (control communities) the implementation was postponed until December 1999.[5] In Figure 1, below, I present the timing of the program.

Figure 1: Timing of the PROGRESA program



There exist a large literature on the evaluation of the average effect of PROGRESA schooling grants. Authors agree in their main conclusions: the program has increased enrollment rates for those children who received the grants, and this positive effect is higher on girls and on children who attend secondary school. Two approaches in this literature can be distinguished according to the methodology applied. Researchers exploited the random assignment of the program at a village level and calculated difference and difference-in-difference estimators. Schultz (2004) is one of the main references. Subsequently, researchers began analyzing ways to improve the effectiveness of the program by estimating structural dynamic models of discrete choice to simulate schooling decisions under alternative policies.[6] Attanasio, Meghir, and Santiago (2005) models schooling as an individual decision and Todd and Wolpin (2006) uses a model of parental decisions about fertility and child schooling.

[5]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.

[6]Eckstein and Wolpin (1989), Rust (1994), and Aguirregabiria and Mira (2007) are exceptional surveys on the estimation of structural dynamic models of discrete choice.

2.3 Summary statistics

The sample used for the estimation of the model includes observations for females from 8 to 17 years old from the October 1998 and November 1999 surveys that provide data for the academic years 98/99 and 99/00.[7] Each wave of the survey includes 9,162 girls belonging to 6,296 families. To identify school dropouts in 1998, I also use information from the October 1997 survey that was collected one year before the implementation of the program. In particular, I use the following question: "Is she attending school now?" In a particular year, a girl is considered a "dropout" if the answer is "no", and a "non-dropout" if the answer is "yes" in the previous year.

By the time of the October 1997 survey, 85% of girls were enrolled in school, 2.2% were working for a salary and 12.8% were neither in school nor working; so I assume they were at home helping in housework. The 15% of girls that were no enrolled in school in 1997 are the 1998 dropouts.

The sample shows two patterns that are usual in net enrollment rates in developing countries: enrollment rates decrease with age and with school level, as it is shown in Figures 2 and 3. Until the age of 11 most of the girls are enrolled in school. Beyond that age enrollment rates decrease more than 10 percentage points for an additional year. There are two grades in which enrollment rates go down remarkably: grade 6, when girls finish primary school, and grade 9, when girls finish secondary school. In both graphs it is evident that girls leave school to stay at home, and not to work for a salary.

[7]Six and seven years old girls are not included because PROGRESA grants are given to those children who have completed at least second grade in primary school. Therefore, a children aged seven or less is not entitled to receive a grant. Additionally, although the entrance in primary school is delayed one or two years, enrollment rates in first and secibd grade in primary school were above 96% in the 1998 survey.

Figure 2: Distribution of choices in academic year 98/99 by age

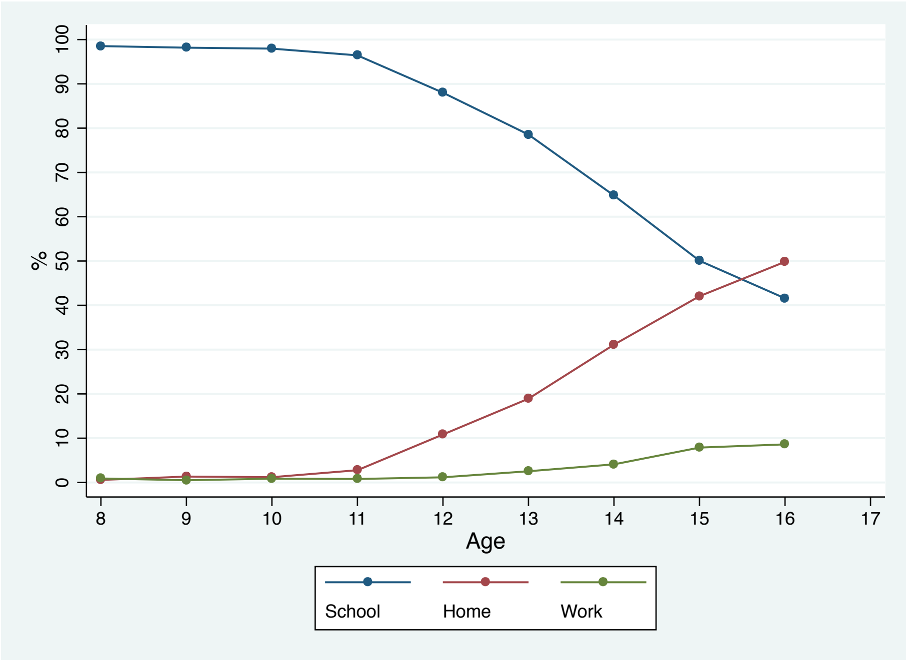
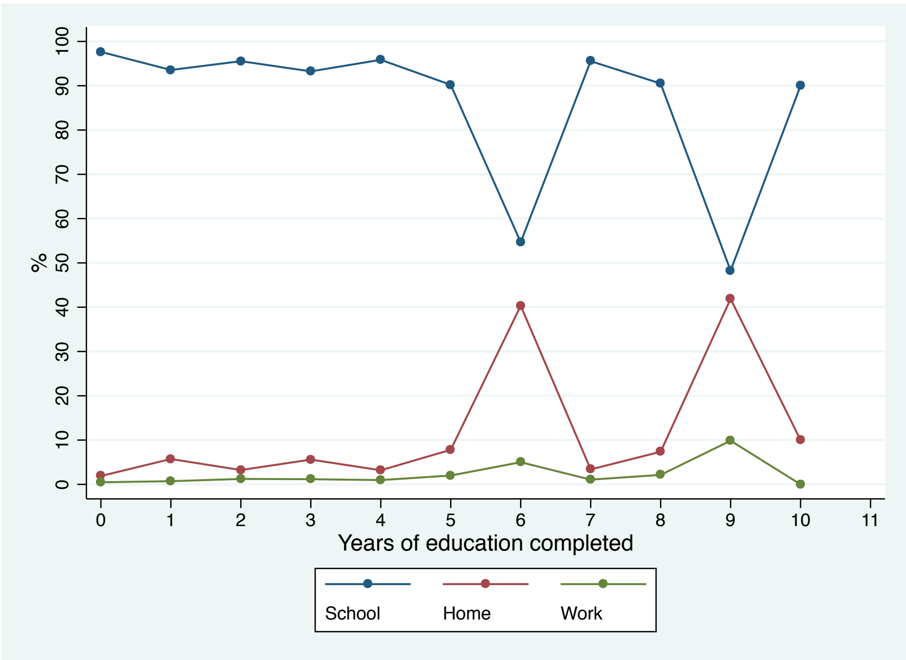


Figure 3: Distribution of choices in academic year 98/99 by stock of education



The enrollment decision, either to go or not to go to school, seems to be highly persistent.

As it is shown in Table 2, the likelihood of going back to school for a girl who was not attending in the previous year and has completed the primary education level goes between 18 and 31%.

Table 2: Choices' transition probabilities by education level (in %)

Incomplete primary school				Completed primary school			
Choice at t	Choice at $t+1$			Choice at t	Choice at $t+1$		
	Enroll	Home	Work		Enroll	Home	Work
Enroll	96.2	2.5	1.3	Enroll	77.8	18.6	3.6
Home	42.2	54.0	3.8	Home	18.0	73.1	8.9
Work	75.3	15.9	8.8	Work	30.6	50.2	19.2

The information contained in the PROGRESA surveys refers to individual characteristics, family composition, parents' activities and background, and community characteristics. Descriptive statistics for selected variables in the academic year 98/99 for non-dropout and dropout girls are presented in Tables 3 and 4.

Table 3: Summary statistics for Non-dropout Girls

Variable	Mean	Std. Dev.	Min.	Max.
Age	11	(2.2)	8	16
Years of education	4.2	(2.1)	0	10
Potential daily wage	19.5	(4.7)	8.5	41.1
Proportion of girls belonging to a poor family	0.9	(0.3)	0	1
Number of adults	3.3	(1.5)	0	12
Number of children	3.7	(2)	0	15
Number of siblings aged 5 or less	1.2	(1.2)	0	11
Proportion of girls whose father is present at home	0.9	(0.3)	0	1
Mother's years of education	2.9	(2.5)	0	18
Proportion of girls that reside in a community with secondary school	0.3	(0.5)	0	1

Table 4: Summary statistics for Dropout Girls

Variable	Mean	Std. Dev.	Min.	Max.
Age	13.3	(2.1)	8	16
Years of education	4.7	(2.2)	0	9
Potential daily wage	23.3	(5.7)	8.4	42.9
Proportion of girls belonging to a poor family	0.9	(0.3)	0	1
Number of adults	3.5	(1.5)	1	10
Number of children	3.8	(2.1)	0	11
Number of siblings aged 5 or less	1.2	(1.2)	0	7
Proportion of girls whose father is present at home	0.9	(0.3)	0	1
Mother's years of education	1.9	(2.1)	0	16
Proportion of girls that reside in a community with secondary school	0.2	(0.4)	0	1

The mean female non-dropout is 11 years old and has four years of education completed. Her mother has completed three years of education. She has four siblings, one of them aged five years old or less. There are three adults individuals living in the same household. If she decides to work she can earn 20 pesos per day, an amount of money higher than the amount of the comparable scholarship scholarship. Of the non-dropouts, 86.5% belong to a poor family, and 7.3% do not live with their fathers. Of the non-dropout, 35% have a secondary school in their village of residence.

The mean female dropout is 13 years old and has five years of education completed. Her mother has completed two years of education. She has four siblings, one of them aged less than six years old. If she decides to work she can earn 23 pesos per day. Of the dropouts, 87.3% belong to a poor family, and 7.5% do not live with their fathers. Of the dropouts, 22% have a secondary school in their village of residence.

Comparing both groups we can conclude that dropouts have less educated mothers, a higher proportion of them have to travel to other communities to attend secondary school, and if they work they receive a higher salary than non-dropouts.

3 Model and Empirical implementation

3.1 The general model

In this section, I present a dynamical behavioral model of schooling decision for girls aged 6 (the official age to enter school) to 17 (the stopping period).[8] At each age t, a girl chooses one of three mutually exclusive actions: work for a salary ($a_{it} = w$), go to school ($a_{it} = e$), or stay at home to help in housework ($a_{it} = h$). This is consistent with assuming that parents make decisions in the best interest of each of their children; therefore, there are no interconnections between the decisions of children that belong to the same family. Let Ω_{it} denote the state vector that contains all variables known by girl i at age t, which have an impact on her current and future choices. Among other components, it also includes the girl's stock of education and the uncertainty about the evolution of her future stock of education. The probability of passing the grade at age t for grade g is denoted by π^s_{tg}, that is, the transition probability for the girl's stock of education. At age 18, girls either work and earn wages in accordance to their levels of education or stay at home.

[8] The reason for choosing 17 as the stopping age is that in the database, all the women aged 18 and above report that they are not enrolled in formal education.

Period t alternatives are chosen to maximize the intertemporal utility function

$$\mathbb{E}[\sum_{j=0}^{T-t} \beta^j u(a_{i,t+j}, \Omega_{i,t+j})|a_{it}, \Omega_{it}] + \beta^{T-t+1}\mathbb{E}[V^{T+1}(\Omega_{T+1})|a_{it}, \Omega_{it}], \tag{1}$$

subject to the evolution of future values of the state variables, particularly to the probability of passing a grade π^s_{tg}. β is the intertemporal discount factor, $V^{T+1}()$ is the terminal value function, $\mathbb{E}_t$ is the expectation operator conditional on the state, and $u(a_{it}, \Omega_{it})$ is the instantaneous utility function at age t for individual i that is specific for each choice a. By Bellman's principle of optimality, the choice specific value functions can be obtained using the recursive expression:

$$v(a, \Omega_{it}) \equiv u(a, \Omega_{it}) + \beta\mathbb{E}[\max_{a\in A} v(a, \Omega_{i,t+1})|a_{it}, \Omega_{it}], \tag{2}$$

for $a = w, e, h$ and $t \leq T-1$, and $v(a, \Omega_{it}) = u(a, \Omega_{it}) + \beta\mathbb{E}[V^{T+1}(\Omega_{T+1})|a_{it}, \Omega_{it}]$ for $a = w, e, h$ and $t = T$. The optimal decision rule is then:

$$\alpha(\Omega_{it}) = \arg\max_{a\in A} v(a, \Omega_{it}). \tag{3}$$

In the database there is information on the individual's action a_{it} and a set of the individual's characteristics X_{it}. From an econometric point of view, the state vector includes two subset of state variables: $\Omega_{it} = (X_{it}, \epsilon_{it}, \mu_i)$. X_{it} are observed variables, and ϵ_{it}, a random time variant shock, and μ_i, a random time invariant variable, are unobserved variables.

3.2 Utilities

The per-period utility function of working is:

$$u(w, \Omega_{it}) = \eta w_{it} + \epsilon^w_{it}, \tag{4}$$

where w_{it} is the potential wage a girl can earn.[9]

The per-period utility function of attending school is:

[9]Because it is reported only in a small percentage of the cases in the survey, it is estimated by ordinary least squares (OLS). For more details see the Appendix.

$$u(e, \Omega_{it}) = \mu_i + \alpha_1 \eta G_{it} + \alpha_2 (1 - \mathbf{1}[a_{i,t-1} = e]) + \alpha_3 S_{it} + \alpha_4 AS_{it} + \alpha_5' x_{it}^e + \epsilon_{it}^e, \tag{5}$$

where μ_i is the unobserved type, individual specific and time-constant. G_{it} is the potential grant amount, that takes a positive value if the child belongs to a poor family, resides in a treated community, and is attending a grade between third year of primary school and third year of junior secondary school. $(1 - \mathbf{1}[a_{i,t-1} = e])$ is the switch cost paid by dropout girls (cost related to changing from not attending to attending school). S_{it} is the girl's stock of education. AS_{it} is a dummy equal to 1 if there is a junior secondary school in the community where the girl resides (is a measure of the direct cost of attending secondary school).

x_{it}^e is a set of individual and family characteristics that includes a dummy variable equal one if the child is behind in school, a dummy variable equal one if the child has graduated from primary school, a dummy variable to reflect the effect of graduation from secondary school, and mother's schooling.

The per-period utility function of staying at home is:

$$u(h, \Omega_{it}) = \delta_1 C5_{it} + \delta_2 SI_{it} + \delta_3 FSize_{it} + \delta_4' x_{it}^h + \epsilon_{it}^h, \tag{6}$$

where $C5_{it}$ is the number of siblings aged five years old or less. SI_{it} is the number of sisters aged from 12 to 16, and $FSize_{it}$ is the number of family members. x_{it}^h is a set of individual and family characteristics that includes the age of the child, and a dummy equal one if the girl's father lives with his family.

3.3 Assumptions

On random shocks: ϵ_{it}^a for $a = w, e, h$ is a random variable that affects the utility of action a in period t for individual i. It is observed by the individual but not by the econometrician. The ϵ_{it}^a's satisfy the conditional independence assumption, that is, they are independent across choices, individuals and periods with distribution $F_\epsilon(.)$.

On utilities: $u(a_{it}, \Omega_{it})$, the utility functions, are additively separable in observables and unobservables:

$$u(a_{it}, \Omega_{it}) = \tilde{u}(a, X_{it}) + \epsilon_{it}^{a} \tag{7}$$

Thus, the optimal decision rule becomes

$$\alpha(X_{it}, \epsilon_{it}) = \arg\max_{a \in A} v(a, X_{it}) + \epsilon_{it}^{a} \tag{8}$$

And, for any $(a, X) \in A \times \mathbb{X}$, the conditional choice probability is:

$$Pr(a|X) = \int \mathbf{1}[v(a, X_{it}) + \epsilon_{it}^{a} > v(a^{'}, X_{it}) + \epsilon_{it}^{a^{'}} \forall a^{'}] dF_{\epsilon}(\epsilon_{it}) \tag{9}$$

On unobserved heterogeneity: Following Heckman and Singer (1984) there are M types of individuals, for M a finite set of types. μ_m is the parameter related to type m and π_m is the proportion of the population of that type.[10] Girls are heterogeneous in their ability at school. Each girl knows her own type but it is not observed by the econometrician.

On transition probabilities: π_{tg}^{s}, the transition probability of the stock of education, is exogenous and does not depend on effort or on the willingness to continue schooling. It varies with the grade and the age of the individual and it is known to the individual.[11] The age of the girl, the amount of the grant, and the amount of the salaries evolve deterministically.[12]. The availability of secondary schools remain constant since the academic year 1999/2000. The girls' mothers stock of education is constant. To control for the socioeconomic situation of the family I use the poor family indicator reported in PROGRESA surveys. This indicator does not vary across time. Girls expect that the composition of her family will not change after 1999. Work status for the girl's mother is assumed time-invariant and identified with her work status reported in 1999. If her father does not live with his family I assume that he is not present at home from 1999 onwards. The number of sisters, brothers, and siblings aged five years old or less evolve with the age of the siblings and I assume that there are no newborn children through

[10]Types probabilities are estimated using a logit model and depend on family composition variables.

[11]It is also different between those girls that receive PROGRESA grants and those who do not receive the aid, because the grant could be an incentive to perform better at school.

[12]The evolution of the amount of the grant from 1998 to 2007 is observed and reported in Oportunidades (2008) The evolution of salaries in the period 1998-2007 is constructed using observed salaries in 1998 and updating them with the annual increase in the minimum wage for Mexico reported in CONASAMI (2008).

all periods.

On individual decision approach: I assume that each girl is a single decision unit. The model presented so far is valid if it is the girl or her altruistic parents who chose the girl's action that maximize her lifetime welfare. It may be difficult to accept that a parent that decides to keep a girl at home is doing an altruistic decision for her/his daughter. This assumption is more reliable if we remember that most of girls aged 18 years old and older have their own families and they are housewife. Their parents may consider that by keeping them at home when they are young they are accumulating housewife skills. In the present model interrelationship of schooling decisions across siblings is not directly considered. The individual decision assumption is relaxed in two ways. First, by allowing the girl's utility of staying at home to vary with several family composition variables. Second, the unobserved type, that enters the utility of attending school, is affected by the number of adults and children in the family and by the girl's birth order. It can be argued that parents make schooling decisions for all their children simultaneously. Particularly, the decision of whether or not to send a daughter to school is affected by the number, ages, gender, and action chosen for the other children in the family. In the model, to choose the action for a girl, parents take into account the number of children they have, their ages, and their genders; parents also consider whether the mother is working outside the household and the total number of adults in the family. The assumption that the decision of one girl in a family does not depend on the decision of other girls in the family may be strong, because, as seen in Table 7 in the Appendix, there is a positive correlation between a girl's decision of attending school and the proportion of sisters that do attend school even after controlling for family composition variables. For girls attending secondary school, the proportion of sisters attending the same level of education only matters, whereas for girls in primary school there is a positive correlation with the proportion of sisters attending both levels of education. Assuming that a girl's participation in primary school does not depend on her brothers choices seems more plausible because there is no correlation between girls' participation in primary school and the proportion of brothers attending school, whereas there is evidence of correlation between girls' participation in secondary school and the proportion of brothers attending secondary school.

3.4 Identification discussion

There are two concerns about the identification of the parameters in the proposed model: state dependence in the utility of attending school and identification of the effect of PROGRESA grants.

State dependence and unobserved heterogeneity: I introduce state dependence in the model in two ways. First, by allowing the utility of attending school to depend on the dropout indicator D_i. Second, state dependence is also present through the stock of education observed in the year before the implementation of the PROGRESA program, $S_{i,97}$, that affects the utility of attending school in 1998. Since we do not know the complete story of schooling decisions before September 1997, having $S_{i,97}$ in the utility of attending school imposes an initial condition problem. Both variables D_i and $S_{i,97}$ are correlated with the unobserved type μ_i. This correlation is explicitly introduced in the model by making the type probability a function of both S_{it} and D_i. The type probability function is:

$$\pi_m = Pr(\mu_i = \mu | S_{i,97}, \mathbf{1}[a_{i,97} = e], Z_{i,97}) = \Phi(\zeta_0 + \zeta_1 S_{i,97} + \zeta_2 \mathbf{1}[a_{i,97} = e] + \zeta_3' Z_{i,97}), \qquad (10)$$

where $Z_{i,97}$ is a set of individual, and family characteristics that includes the birth order of the child, and mother's educational level. note, that although the estimated parameters in the type probability function are highly significant the identification of the unobserved heterogeneity using a panel with only two waves of data in a model that includes state dependence is tenuous and relies strongly on functional forms.

Grant effect: The effect of the grant in the utility of attending school is modeled as a proportion of the impact of the wage in the utility of working. The model can then reflect a different effect given by one peso received as a grant or one peso received as a salary on the decision of attending school. For the identification of both effects it is necessary to have two different sources of exogenous variation. Wages vary with the girls' age, their stock of education, and a set of labor market variables of their village of residence. The amount of the grant also varies with the girls' age and their stock of education, and, most importantly, it has an exogenous (random) variation between girls who reside in treatment and control communities.

3.5 Likelihood

Define $\theta^a = \{\eta, \alpha_1, ..., \alpha_5, \delta_1, ..., \delta_4', \lambda\}$ as the set of parameters in utilities and in the terminal value, and $\theta^m = \{\zeta_0, ..., \zeta_3'\}$ as the set of parameters in the type probability function. Let $\rho = \{\theta^a, \theta^m, \{\mu_m\}_{m=1}^M, \beta, \pi_{tg}^s\}$, the set that includes all the parameters to be estimated in the model and the transition probability of the stock of education. Suppose $\tilde{u}(a, X_{it}), V^{T+1}()$ and $F_\epsilon()$ are known up to ρ. A girl's contribution to the likelihood is:

$$
\begin{aligned}
l_i(\rho) = \sum_m^M \Big\{ \prod_k \Big[\sum_a \mathbf{1}(a_{ik} = a) \times Pr(a_{ik} = a | X_{it}, S_{it}, \mathbf{1}[a_{i,k-1} = e], \mu_m, \theta^a, \pi_{tg}^s, \beta) \Big] \\
\times Pr(\mu_m = \mu | S_{i,97}, \mathbf{1}[a_{i,97} = e], Z_{i,97}) \Big\},
\end{aligned} \tag{11}
$$

where $m = 1, ..., M$ indexes the types of individuals, $k = 1998, 1999$ indexes the waves of the survey, $a = w, e, h$ indexes the three choices, and $t = 8, ..., 17$ indexes the age of the girl. The sample log-likelihood is then $L(\rho) = \sum_i \ln l_i(\rho)$.

To evaluate the l_i for a particular value of ρ it is necessary to know the optimal decision rules $\alpha(X_{it}, \epsilon_{it}, \rho)$. Therefore, for each trial value of ρ the value functions $v(a, \Omega_{it})$ have to be computed. The expression for the value functions at subsequent ages is computed recursively starting from age 18 and working backwards until the current age t. Under the assumption that the unobserved state variables ϵ_{it}^a are drawn from an extreme value distribution, conditional choice probabilities and recursive value functions in Eq. 2 have convenient (logistic) closed forms.[13] The model is estimated by a combination of maximum likelihood for $\theta^a, \theta^m, \{\mu_m\}_{m=1}^M$, and a grid search for the discount factor β.

4 Results

4.1 Parameter estimates

Maximum likelihood estimates of the model's structural parameters are reported in Tables 8, and 9 in the Appendix. The probabilities of passing grade s at age t, π_{tg}^s, used in the estimation

[13] See the Appendix for the explicit functional form of value functions, conditional choice probabilities and $\mathbb{E}max$ function.

of the model are reported in Tables 10 and 11 in the Appendix.

The estimated parameters in the three instantaneous utilities have the expected signs and the most relevant parameters for the policy analysis pursued in this paper are statistically significant. The utility of attending school is higher for more educated girls, who have more educated mothers, and living in communities where there is a secondary school. Salaries have a positive effect on the utility of working. The utility of staying at home is higher for older girls, belonging to a family with at least one children aged five years old or less.

The estimation of the model has been done so far with two types of individuals, that are well identified. The high type individuals have a higher utility of attending school with an estimated unobserved effect equal to 16.2. The corresponding estimated value for low types is 12. The probability of being of high type is higher for girls with a higher stock of education, that were attending school in 1997, and whose mothers have achieved a higher educational level. The estimated coefficient of $\mathbf{1}[a_{i,97} = e]$ is highly significant, that is, the model relates the types with the enrollment status in the initial period.

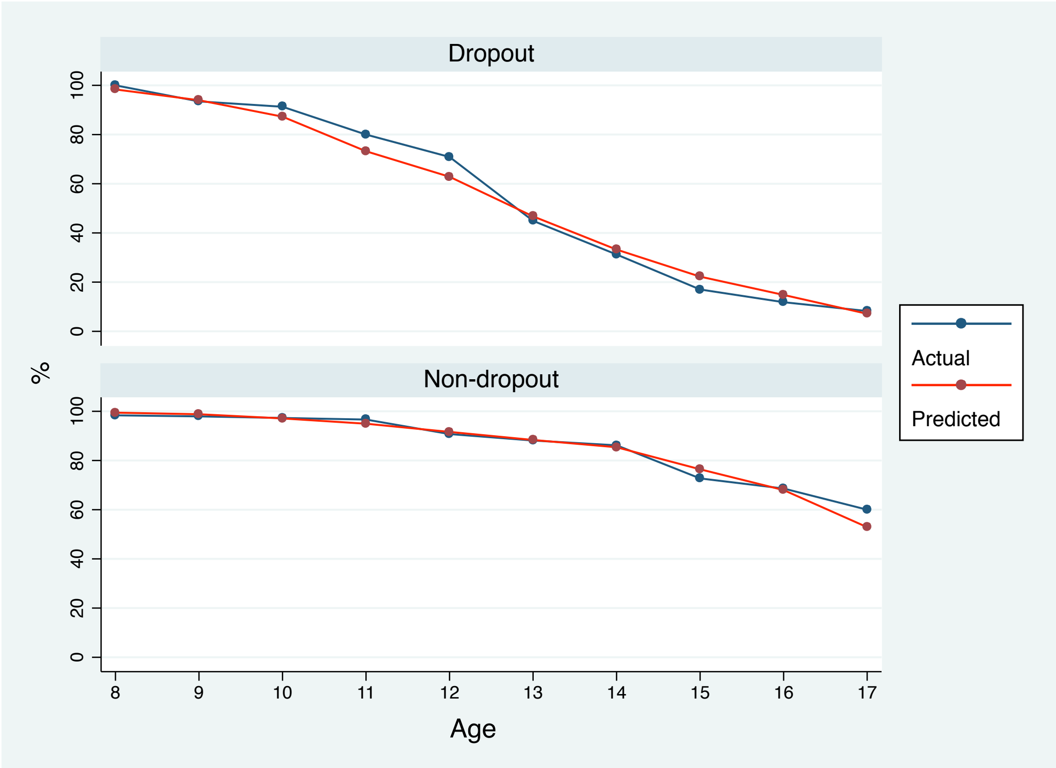
4.2 Model Validity

The validity of the structural parameters estimates relies strongly on the functional form assumptions made on utilities, the value function, and the type probability. Thus, it is crucial to test the validity of the estimated model. In what follows, I present evidence on the validity of the estimated parameters.

First, I compare the distribution of predicted choice probabilities obtained with the estimated parameters with the actual choices the individuals in the sample have made.[14] As seen in Figure 4, the model does quite well in predicting distribution of choices by ages for non-dropouts, and it fits the distribution of choices of dropouts reasonable well, although it is not as accurate as in the case of non-dropouts.

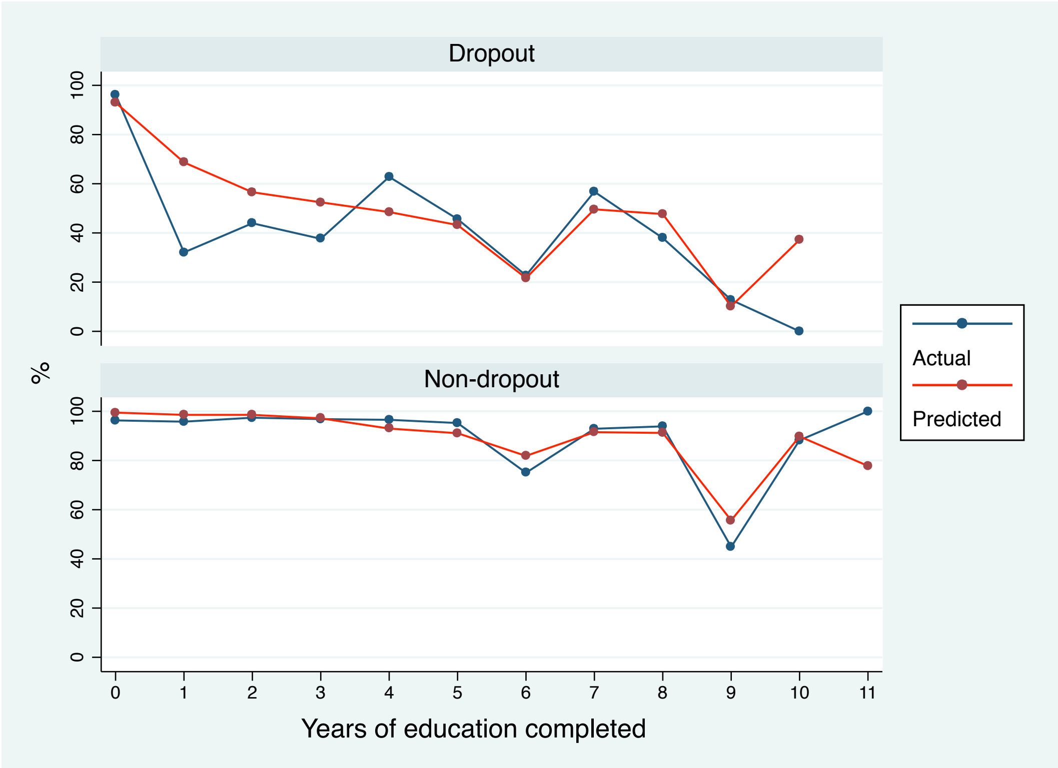
[14]Predicted conditional choice probabilities are computed following Carro and Mira (2006). The procedure is explained in the Appendix.

Figure 4: Actual and predicted enrollment rates by age (%)



Predicted choice probabilities by years of schooling completed, reported in Figure 5, reflect the main patterns in the actual distribution of choices: enrollment rates for non-dropouts are always higher than for dropouts; enrollment rates for both groups decrease as the stock of education increases; the lowest enrollment rate in primary school appears in the last grade, that is grade six; and the lowest enrollment rate considering all grades come in the last year of junior secondary school. note that the fit of the model is better for non-dropouts than for dropouts.

Figure 5: Actual and predicted enrollment rates by stock of education (%)



Second, I compare the estimated grant effects with the difference-in-difference estimates reported in Schultz (2004). The effect of PROGRESA grants in the present model is computed by comparing the choices predicted when girls are receiving the grant with the choices predicted when the grant amount is set to zero for all girls. The results obtained in this paper agree with those reported elsewhere in the literature, suggesting that the model does well in fitting the effects of PROGRESA grants in average, as it is reported in Table 5.

Given the evidence presented so far, I do think that the estimated model fits the data reasonable well, so in what follows I continue with the simulation of enrollment and reentry rates under the presence of several policies.

Table 5: Comparison of the estimated effect of the grant using diff-in-diff vs. structural estimation.

	Educational level		
	Primary	Secondary	
		6th grade	Above 6th grade
Grant effect			
Diff-in-diff	2.93	10.0	4.48
Structural model	2.43	9.45	6.19
Estimated enrollment without interventions			
All	90.4	49.5	69.2
Dropout	52.9	13.2	18.2
Non-dropout	93.7	72.4	76.0

4.3 Counterfactual analysis

Although PROGRESA grants do not sustantially increase school reentry rates among girls, perhaps other policies do. In what follows, I analyze the effectiveness of several policies in increasing reentry rates by means of counterfactual exercises. Results are presented in Table 6 and details on the policies follow bellow. It is worth to note that counterfactuals exercises obtained by using the parameters related with PROGRESA grants are more credible than other exercises. The reason is that the identification of PROGRESA grants parameters is obtained with the exogenous variation provided by the random assignment of the scholarships, whereas there is no natural experiment behind the identification of the other policy related parameters in the model.

Table 6: Increase in school attendance due to different policies (in %)

	Non-dropout			Drop-out		
	Primary	Secondary		Primary	Secondary	
		Grade 6	Grade 7 and more		Grade 6	Grade 7 and more
Baseline enrollment rates (estimated without policy)	94	72	76	53	13	18
Policy:						
PROGRESA grant	1.1	10.2	2.3	3.1	8.3	6.4
Duplicate PROGRESA grant	1.9	17.6	3.8	6.3	19.8	13
Free access to daycare center	2.9	19.4	6.5	11.7	20.8	8.5
Availability of secondary school in almost all villages	0.95	4.3	3.6	4.0	2.8	3.4
Sample size in 1998	5,596	1,122	1,107	550	698	89

Duplicate the amount of PROGRESA grants in secondary school: It may be the case that the low effectiveness of PROGRESA grants in increasing school attendance among dropouts is related with the design of the scholarship. In particular, the amount of the grants may be low to obtain the desired results. One alternative is to set the value of the grant at the level of the wages the girls can earn in the labor market. This change in PROGRESA grants means that the amount of the grant needs to be duplicated, a financial effort that the government may not be able to afford.

The results show that for non-dropouts increasing the amount of the grant produces a less than proportional increase in enrollment rates, while duplicating the grant more than duplicates the effect on dropouts. But still, even using a grant scheme that sets the amount of the grant equal to the labor salary, enrollment rates will be bellow 35%. So other policy interventions that may generate further increases in enrollment rates among dropouts should be explored.

Community nurseries/kindergartens: Suppose all the children aged five years old or less in the family are sent to a (free) daycare center.[15] Girls may be no longer needed at home to look after them and may go back to school. In the model, the utility of staying at home is (positively) related with the number of children aged five or less years old in the family. The effect of this policy on girls' school enrollment can be approximated by simulating girls' choices after setting

[15] I do not discuss how the daycare center would be financed.

the number of children aged five or less years old equal to zero, while keeping constant the total number of family members. Note that imposing that the number of children less than six in the family is equal zero may overestimate the actual effect of having free access to a daycare center. Not having children below six at home means that the girl will not have to take care of them throughout the day, although having access to a child care facility will reduce the need of help in taking care of the youngest children in the family for at most eight hours. In the case that school hours fully coincide with the time the children can be left in the daycare center, the upward bias in the estimation will be reduced. The approximate effect of availability of nurseries is much more highter than the effect of PROGRESA grants, and even higher than the effect of a grant that sets the amount at the level of local salaries.

Availability of secondary school in almost all villages: As it is shown in Tables 3 and 4, only in 35% of the villages where non-dropouts reside and in 22% of the villages where dropouts reside there exists a secondary school. Not having a secondary school available in the village of residence implies transportation and time costs. Both costs decrease the utility of attending school. If the government establishes a secondary school in at least all villages where the demand is high enough, a positive effect on school enrollment and reentry rates could be expected. I simulate girls' choices by setting the indicator variable of availability of secondary school for girls who reside in villages where the potential number of secondary school students is higher than 25 equal to one. The result is promising for all groups, non-dropouts and dropouts attending primary and secondary school. In primary school, the enrollment rate increases 1 percentage point and the reentry rate increases 4 percentage points, whereas the figures in the first year of secondary school are 4.3 percentage points increment in the enrollment rate and 2.8 percentage points increment in the reentry rate.

5 Conclusions

In this paper I present a dynamic behavioral model of school choices for girls in poor families and estimate its structural parameters using the Mexican PROGRESA database. The estimated structural model fits girl's schooling choices reasonably well. It is able to replicate patterns observed in the actual distribution of schooling choices, and it also matches differences in the distribution of schooling choices between girls who attend school and girls who dropped out. The model explains these differences highlighting the relevance of persistence in the decision of attending school and the importance of the girl's family composition. Results also suggest that unobserved heterogeneity in the utility derived from attending school does explain the observed differences in the decisions taken by female non-dropouts and dropouts.

The evaluation of PROGRESA grants resulting from the estimated model is consistent with previous literature. Grants are a good incentive to keep girls at school but the ones that have dropped out of school return in low proportions. Simulations suggest that cash transfers do not substantially increase school reentry rates, but these economic incentives are successful when the amount of the scholarship is duplicated. As dropouts are mainly at home helping to take care of the youngest children in the family, the availability of daycare centers has a sizable effect on increasing reentry rates and on preventing dropouts by increasing enrollment rates. Enrollment rates for both group targets are also efficiently increased by reducing transportation and time costs in secondary school.

The relevance of family characteristics in girls' schooling choices suggested by the present model, invites future research. The most natural extension is the study of school reentry decisions in the context of a family decision model. The estimation of a model of family child schooling and fertility decisions, similar to the model presented in Todd and Wolpin (2006), allows relaxing the assumption that there is no newborn children in the girls' families.

As a further step, it would be interesting to estimate a collective decision model in which parents make labor and consumption decisions along with schooling decisions for their children. Such a model would allow the analyses of interrelations between parents' labor participation decisions and girls schooling choices in poor families. It can be expected that a girl whose mother works in the labor market would be more valuable at home when the decision is taking to maximize the total utility of the family, rather than maximazing the individual utility of the girl. However, worker mothers in the sample have fewer children than mothers who stay at home, showing that there is some interaction between family composition and parents' labor status. A collective decision model in which parents simultaneously made fertility choices, and decide their labor status and their children schooling would shed light to the relation between both decisions. In the framework proposed, it would be possible to analyze the effect of policies intended to increase children school participation on parents' labor participation and girls schooling.

References

AGUIRREGABIRIA, V., AND P. MIRA (2007): "Dynamic discrete choice structural models: A survey," *University of Toronto. Department of Economics*, Working Paper 297.

ALBARRAN, P., AND O. ATTANASIO (2002): "Empirical implications of limited commitment. Evidence from Mexican villages," *CEMFI and UCL manuscript.*

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., P. SENGUPTA, AND P. TODD (2001): "Progressing Through PROGRESA: An Impact Assessment of a School Subsidy Experiment," *International Food Policy Research Institute, Washington, D.C. Research Report.*

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.*

CARRO, J. M., AND P. MIRA (2006): "A dynamic model of contraceptive choice of Spanish couples," *Journal of Applied Econometrics*, 21, 955–980.

CONASAMI (2008): "Tabla de salarios mnimos generales y profesionales," *Comisin Nacional de los Salarios Mnimos. Secretaria del Trabajo y Previsin Social de Mexico downloadable from www.conasami.gob.mx.*

ECKSTEIN, Z., AND K. WOLPIN (1989): "The specification and estimation of dynamic stochastic discrete choice models: a survey," *Journal of Human Resources*, 24, 562–598.

HECKMAN, J., AND B. SINGER (1984): "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data," *Econometrica*, 52, 271–320.

KEANE, M., AND K. WOLPIN (1997): "The career decisions of young men," *Journal of Political Economy*, 105, 473–522.

Miller, R. (1984): "Job matching and occupational choice," *Journal of Political Economy*, 92, 1086–1120.

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

——— (2008): "Histrico de apoyos monetarios," *Secretaria de Desarrollo Social del Gobierno de Mexico downloadable from www.oportunidades.gob.mx.*

Rubalcava, L., and D. Thomas (2000): "Family Bargaining and Welfare," *RAND - Labor and Population Program - Paper Series*, 00-10.

Rust, J. (1994): "Structural estimation of Markov decision processes," *Handbook of Econometrics*, Vol.1A, McFaden D, Engel R (eds).

Schultz, T. 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 (2006): "Assessing the Impact of a School Subsidy Program in Mexico: Using a Social Experiment to Validate a Dynamic Behavioral Model of Child Schooling and Fertility," *American Economic Review*, 96-5, 1384 – 1417.

UNESCO (2007): "State of education report," *UNESCO Institute for Statistics.*

Valdes, M. N. (2008): "Did PROGRESA send drop-outs back to school?," *Universidad Carlos III de Madrid. Department of Economics. Working Paper 08-59*, Economic Series (26).

Appendix

Value functions

The value function for choosing to attend school is:

$$\begin{aligned} v(e, X_{it}) = {} & \tilde{u}(e, X_{it}) \\ & + \beta\pi^{s}_{tg}\mathbb{E}_{\epsilon}[\max_{a\in A}\{v(a, X_{i,t+1}) + \epsilon^{a}_{it}\}|X_{it}, S_{i,t+1} = S_{it} + 1, a_{it} = e] \\ & + \beta(1 - \pi^{s}_{tg})\mathbb{E}_{\epsilon}[\max_{a\in A}\{v(a, X_{i,t+1}) + \epsilon^{a}_{it}\}|X_{it}, S_{i,t+1} = S_{it}, a_{it} = e], \end{aligned}$$

for $a = w, e, h$ and $t \leq T - 1$. At age $t = T \equiv 17$ it is:

$$\begin{aligned} v(e, X_{T}) = {} & \tilde{u}(s, X_{T}) \\ & + \beta\pi^{s}_{tg}V^{T+1}(X_{T+1}, S_{i,T+1} = S_{iT} + 1) \\ & + \beta(1 - \pi^{s}_{tg})V^{T+1}(X_{T+1}, S_{i,T+1} = S_{iT}). \end{aligned}$$

The value function for working (or staying at home) is:

$$\begin{aligned} v(w, X_{it}) = {} & \tilde{u}(w, X_{it}) \\ & + \beta\mathbb{E}_{\epsilon}[\max_{a\in A}\{v(a, X_{i,t+1}) + \epsilon^{a}_{it}\}|X_{it}, S_{i,t+1} = S_{it}, a_{it} = w], \end{aligned}$$

for $a = w, e, h$ and $t \leq T - 1$. At age $t = T \equiv 17$ it is:

$$v(w, X_{T}) = \tilde{u}(w, X_{T}) + \beta V^{T+1}(X_{T+1}, S_{i,T+1} = S_{iT}).$$

I assume that girls do not attend school beyond 18 years old. At this age they receive a terminal value that depends on their stock of education and on the potential salary they could earn at the local labor market. The terminal value function is:

$$V^{T+1} = \lambda S_{i,18}.$$

In all cases below, $\mathbb{E}max$ function are as follows:

$$\mathbb{E}_\epsilon[\max_{a\in A}\{v(a, X_{i,t+1}) + \epsilon_{it}^a\}|X_{it}, S_{i,t+1}, a_{it}] = ln(\sum_{a\in A} exp(v(a, X_{t+1}))) + E,$$

where E is the Euler constant (0.577215665). This expression is given by the extreme value distribution and by the conditional independence assumptions on ϵ_{it}^a.

Conditional choice probabilities

Assuming the ϵ_{it}^a are drown from an extreme value distribution and are conditional independent, the probability of choosing action a at time t is:

$$Pr(a_{it} = a'|X_{it}) = \frac{\exp v(a', X_{it})}{\sum_{a\in A} \exp v(a, X_{it})}$$

Predicted probabilities

Following Carro and Mira (2006), predicted conditional choice probabilities for each girl are computed as the weighted average of conditional choice probabilities for each unobserved type, with weights given by the ex post probability that the girl is of each type conditional on her stock of education and choice in period k=1998,1999.

$$\mathbb{P}_{iak} = \sum_{m}^{M} \mathbb{P}_{iamk}\mathbb{P}(\mu_i|A_i, S_{i,97}, D_i)$$

$$\mathbb{P}(\mu_i|A_i, S_{i,97}, D_i) = \frac{\mathbb{P}(\mu_m, A_i|S_{i,97}, D_i)}{\mathbb{P}(A_i|S_{i,97}, D_i)}$$

$$\mathbb{P}(\mu_m, A_i|S_{i,97}, D_i) = \mathbb{P}(A_i|S_{i,97}, D_i, \mu_m)\mathbb{P}(\mu_m|S_{i,97}, D_i)$$

$$\mathbb{P}(A_i|S_{i,97}, D_i, \mu_m) = \sum_{k=98}^{99} \sum_{a} \mathbf{1}[a_{ik} = a]\mathbb{P}_{iamk}$$

$$\mathbb{P}(A_i|S_{i,97}, D_i) = \sum_{m}^{M} \mathbb{P}(a_i|S_{i,97}, D_i, \mu_m)\mathbb{P}(\mu_m|S_{i,97}, D_i),$$

where $A_i = a_{i,98}, a_{i,99}$. $\mathbb{P}_{iamk}$ and $\mathbb{P}(\mu_m | S_{i,97}, D_i)$ are obtained from the model given parameter estimates.

Correlation in school attendance among siblings

Table 7 bellow shows the OLS estimates of the coefficients of the following simple linear model:

$$y_i = \beta_0 + \beta_1^{'} X_i,$$

where y_i is equal one if the girl is attending school, and X_i includes the age of the girl (*age*), the proportion of sisters aged below 12 attending school (*asgypr*), the proportion of sisters aged 12 or more years old attending school (*asgopr*), similar variables for the proportion of brothers attending school (*asbypr* and *asbopr*), a dummy equal one if the girl's father is living with her family (*f_hogar*), number of brothers between 6 and 16 years old (*boy*), number of sisters between 6 and 11 years old (*girl*11), number of sisters between 12 and 16 years old (*girl*16), number of sisters between 17 and 18 years old (*girl*16), number of children aged less than 5 years old (*baby*), a dummy equal one if the girl's mother works for a salary (*work_m*), and an indicator of the socioeconomic status of the girl's family (*poor*).

Table 7: Correlation in school attendance among siblings. Output variable: girl is attending school.

Girls in primary school			Girls in secondary school		
Variable	**Coefficient**	**(Std. Err.)**	**Variable**	**Coefficient**	**(Std. Err.)**
age	-0.043***	(0.002)	age	-0.041***	(0.005)
asgypr	0.701***	(0.041)	asgypr	0.007	(0.052)
asgopr	0.137***	(0.010)	asgopr	0.959***	(0.017)
asbypr	0.046	(0.039)	asbypr	0.014	(0.055)
asbopr	-0.005	(0.010)	asbopr	0.038**	(0.016)
f_hogar	-0.021	(0.023)	f_hogar	-0.009	(0.037)
boy	-0.002	(0.006)	boy	-0.010	(0.009)
girl11	0.017**	(0.007)	girl11	0.000	(0.013)
girl16	-0.012	(0.008)	girl16	-0.004	(0.011)
girl18	0.017*	(0.010)	girl18	0.068***	(0.015)
baby	-0.003	(0.004)	baby	-0.007	(0.006)
work_m	-0.008	(0.019)	work_m	0.030	(0.029)
poor	0.010	(0.018)	poor	0.021	(0.023)
Intercept	0.575***	(0.061)	Intercept	0.576***	(0.110)

Significance levels : $*$: 10% $**$: 5% $***$: 1%

Estimation of salaries

The salary for a girl i residing in village l that chooses to work is computed using the OLS parameters of the following equation:

$$ln(w_{il}) = \gamma_0 + \gamma_1 ln(w_l) + \gamma_2 S_i + \gamma_3 age_i + \gamma_4 distmetro_l + \gamma_5 p_l + \omega_{il},$$

where w_l is the agricultural wage in community l, $distmetro_l$ is the distance (km) from the community where the girl resides to the nearest metropolitan area, and p_l is a dummy equal one if the village l was a treatment locality.[16]

A sample selection problem arises in the estimation of the previous equation. The resulting estimated salaries may be not a good approximation of the potential salaries for girls who do not work and for girls who do work but do not report their salaries. This problem is solved by adding the assumption that the transitory shocks to potential earnings ω_{il} are not observed before the girl make her choice. Moreover, the variance of gilrs's salaries may be low since they are expected to work in low skilled homogenous agricultural activities.

[16]Including the dummy p in the equation allows to control for potential general equilibrium effects of the program on the local labor market. Specifically, the introduction of the program may reduce the local labor supply of children in the short run inducing an increase in salaries at least the first year of implementation of the grants. This dummy is included in the equation of salaries in 1998, but not in 1999, because after two years the total local labor supply (adults and children) may return to its initial level if some workers move between localities. In localities that are geographically close, those movements of workers are highly possible.

Estimation results

Table 8: Estimates of structural parameters: Instantaneous Utilities and terminal value function

Variable	**Estimate**	**Standard Error**
Schooling utility		
grant	0.06	0.025
switch cost	-1.73	0.072
stock of education	-11.24	0.46
graduation from primary	3.01	0.16
graduation from secondary	4.02	0.51
behind in school (repeater)	-0.10	0.07
existence of secondary school	0.31	0.06
mother stock of education	1.57	0.23
Working utility		
wage	0.03	0.013
Staying at home utility		
age	2.80	0.18
number of children aged 5 or less	2.32	0.28
number of girls aged 12 to 16	-2.92	0.33
number of family members	0.51	0.12
father present at home	-0.15	0.095
Terminal value function		
school	10.46	0.63

Log-likelihood = -6,786.38, Discount Factor = 0.95

Table 9: Estimates of structural parameters: Types and Types probabilities

	Estimate	**Standard Error**
Unobserved type effect		
Type 1	16.18	0.47
Type 2	12.04	0.41
Variables in types probabilities		
stock of education in 1997	12.87	1.66
enrolled in school in 1997	3.77	0.31
birth order	-2.21	0.89
mothers' stock of education	2.22	1.01
constant	-4.15	0.59

Reference category is Type 2

Table 10: Probability of passing grade s at age t for girls who fulfill the conditions required to receive the grant

	Grade (s)														
Age (t)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
6	0.81	0.48	1.00	1.00	1.00	1.00	0.70	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
7	0.93	0.84	0.87	1.00	1.00	1.00	0.65	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
8	0.89	0.82	0.91	0.88	0.80	1.00	0.61	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
9	1.00	0.74	0.87	0.91	0.87	1.00	0.56	1.00	1.00	0.90	1.00	1.00	1.00	1.00	1.00
10	0.93	0.78	0.91	0.88	0.94	0.91	0.00	0.99	0.95	0.79	1.00	1.00	1.00	1.00	1.00
11	0.91	0.73	0.70	0.82	0.89	0.89	0.67	0.95	0.87	0.68	1.00	1.00	1.00	1.00	1.00
12	0.83	0.40	0.85	0.72	0.83	0.88	0.71	0.95	0.78	0.57	1.00	1.00	1.00	1.00	1.00
13	1.00	0.33	0.50	0.65	0.88	0.84	0.56	0.90	0.50	0.46	1.00	1.00	1.00	1.00	1.00
14	1.00	0.25	0.33	0.27	0.73	0.77	0.37	0.80	0.64	0.38	1.00	1.00	1.00	1.00	1.00
15	0.99	0.29	0.00	0.18	0.20	0.61	0.22	0.80	0.71	0.17	1.00	1.00	1.00	1.00	1.00
16	1.00	0.25	0.00	0.00	0.25	0.44	0.19	0.83	0.79	0.19	0.00	1.00	0.00	0.00	0.00
17	1.00	0.20	0.03	0.03	0.32	0.47	0.00	0.76	0.00	0.00	0.00	1.00	0.00	0.00	0.00

Table 11: Probability of passing grade s at age t for girls who do not fulfill the conditions required to receive the grant

	Grade (s)														
Age (t)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
6	0.84	0.51	0.33	1.00	1.00	1.00	1.00	1.00	0.00	0.66	0.00	1.00	0.00	0.00	0.00
7	0.89	0.85	0.87	0.67	1.00	1.00	1.00	1.00	0.00	0.64	0.00	1.00	0.00	0.00	0.00
8	0.93	0.83	0.92	0.88	0.75	1.00	1.00	1.00	0.00	0.63	0.00	1.00	0.00	0.00	0.00
9	0.88	0.77	0.83	0.91	0.82	1.00	0.90	1.00	0.00	0.61	0.00	1.00	0.00	0.00	0.00
10	1.00	0.77	0.77	0.85	0.95	0.91	0.78	0.99	0.00	0.59	0.00	1.00	0.00	0.00	0.00
11	1.00	0.63	0.79	0.74	0.90	0.92	0.67	0.95	0.12	0.58	0.00	1.00	0.00	0.00	0.00
12	0.83	0.60	0.55	0.61	0.86	0.95	0.60	0.90	0.00	0.56	0.00	1.00	0.00	0.00	0.00
13	1.00	0.50	0.45	0.38	0.82	0.82	0.46	0.91	0.65	1.00	0.00	1.00	0.00	0.00	0.00
14	0.50	1.00	0.00	0.36	0.53	0.67	0.14	0.76	0.75	0.24	0.00	1.00	0.00	0.00	0.00
15	0.51	0.33	0.00	0.05	0.53	0.42	0.15	0.68	0.77	0.25	0.40	1.00	0.40	0.40	0.40
16	0.00	1.00	0.00	0.00	0.00	0.17	0.07	0.79	0.77	0.07	1.00	1.00	1.00	1.00	1.00
17	0.38	0.72	0.02	0.00	0.28	0.19	0.00	0.67	1.00	1.00	1.00	1.00	1.00	1.00	1.00