Using a Structural Model of Educational Choice to Improve Program Efficiency

by

Alain de Janvry, Frederico Finan, and Elisabeth Sadoulet

University of California at Berkeley

February 2005

Address of corresponding author: Elisabeth Sadoulet, Giannini Hall 207, University of California, Berkeley, California 94720. Tel: 510-642-7225. Fax: 510-643-8911. E-mail: Sadoulet@are.berkeley.edu

Using a Structural Model of Educational Choice
to Improve Program Efficiency

Abstract

Constructing structural models of educational choice allows to explore design features for educational programs and to predict how the program would perform in alternative contexts, for instance when accompanied by new complementary programs. We use the experience of Progresa, Mexico's ambitious conditional cash transfer program for education in poor rural communities, to construct such a model. The impact of transfers on decisions to enroll in secondary school and to repeat a grade in case of failure is accurately measured due to randomized treatment in a subset of communities. While impact measurements of Progresa on educational attainment are available from reduced form estimates, the structural model allows to decompose the channels of influence in decision making and to measure their relative importance on observed outcomes. We measure the gains from a design where future transfers can be credibly committed in spite of political cycles, and from complementary supply-side programs providing improved off-school support to students and access to better information about job opportunities outside the community offered by education.

1. Using structural models for program design

Designing new social programs and fine-tuning them for maximum impact can be extraordinarily costly and time consuming. As a result, what we typically see being done is putting into place expensive programs that have not been tested for alternative designs, either at a pilot stage or in early years of implementation. Even if randomization is used in implementation, allowing to rigorously identify impact, results only apply to the program as it was implemented and to the specific context where it was applied.

Information available on one of the most innovative and well designed conditional cash transfer (CCT) programs for education, Progresa in Mexico, is of this type. Randomized treatment on a subset of villages allowed rigorous measurement of impact (Parker and Skoufias, 2000; Behrman, Sengupta, and Todd, 2001; Schultz, 2004), but only for the single design that was implemented and for the context that prevailed. While measured impacts were encouraging, they do not provide any leads to help revise the program's design in an attempt at reaching greater efficiency and to anticipate how the program would perform in alternative contexts that could be modified by complementary interventions.

For this reason, in the same way as plant breeders have modeled plant growth to reduce the cost of experimenting with alternative designs in alternative contexts, social scientists have started to develop dynamic structural models of behavioral response to social programs. In the case of Progresa, attempts at structural modeling have been made by Todd and Wolpin (2003) and by Attanasio, Meghir, and Santiago (2002). For the first, the objectives were to predict responses to variations in program design, specifically in the way cash transfers are offered: different amounts, offer of a bonus upon graduation, and making transfers non-conditional on school attendance. For the second, the objective was to see the impact of a revenue neutral change in the structure of the payments in favor of secondary school.

In this paper, we construct a structural model of enrollment decisions in secondary school and of decisions to repeat a grade in case of failure, and of the impact that Progresa's CCTs have on these two decisions. The model is estimated using the randomized treatment to identify the role of CCT. The structural model estimated allows us to derive results from the Progresa experiment that go beyond what could be done with a reduced form estimate of behavioral response. Specifically, we use the model for two purposes.

The first is to decompose the impact of CCT on schooling decisions into four channels of influence: (i) the impact on enrollment of current transfers as they affect the utility for schooling; (ii) the impact on performance at school (probability of making the grade) of Progresa conditionalities regarding school attendance to qualify for the transfers, (iii) the impact of transfers on the decision to repeat the grade in case of failure, and (iv) the impact on current enrollment of expected future transfers attached to continued enrollment. The last two effects allow us to measure the role of an important feature which is always at risk in public programs whose existence depends on the political cycle: the impact on current schooling decisions of the ability to commit that CCT will be maintained beyond the current period if children continue to enroll. Indeed, a major achievement of Progresa is that it survived intact from one presidential cycle to the next (Székely, 2004). Since parents' schooling decisions this year are based on the benefits of both current and expected future transfers in subsequent years, the structural model allows us to measure how much does the certainty of future conditional transfers buys in terms of additional school enrollment and grade repetitions this year.

The second is to simulate the role of complementary supply-side interventions such as improved off-school support to students and access to better information about job opportunities outside the community offered by education. These complementary programs have an effect on non-beneficiaries of CCT, and also an effect on beneficiaries additional to the CCT. Exploring complementarities between demand and supply-side instruments would be difficult to achieve experimentally, but can be explored through simulation in the structural model constructed here.

2. What factors motivate school enrollment and performance?

A structural model of educational attainment must capture the main factors that motivate children to enroll in school and that help them succeed in school. There are three.

The first factor is the current utility to the family that is provided by sending children to school, for instance because educated parents prefer educated children. Particular children also have differential preferences for attending school (Akerlof and Kranton, 2002). Many empirical studies have related a child's school achievement to his own and his parents' characteristics (Haveman and Wolfe, 1995; Acemoglu and Pischke, 2001). If viewed as a "consumption good", schooling is purchased according to preferences subject to a budget constraint. Increasing school attendance could thus be achieved by campaigns to alter parents' and children's preferences, by direct transfers to relax their budget constraint, or by conditional transfers that reduce the cost of schooling.

The second factor that affects schooling decisions is that education can be viewed as an investment in future earnings, as documented in the work of Willis and Rosen (1979). In this case, attendance should be influenced by the expected return to schooling and, if the capital market works, it is only to the extent that children's and parents' characteristics influence the return to schooling that they should affect attendance to school in a given context. As a first approximation, school attendance would mostly be increased by the development of an active labor market that rewards education. In a world of imperfect capital markets, the household's characteristics would also affect schooling decisions through the shadow price of capital. As the development of a long term credit market for investment in primary and secondary schooling is not easy to implement, transfers or conditional transfers can help households close the gap with their optimal investment. Note, however, that it is the perceived rather than the actual market return to schooling that dictates schooling decisions, and hence imperfect information on job market opportunities may substantially reduce the incentives for a child to get a higher educational level. This suggests that promoting access to, or better information about, job opportunities may encourage a higher demand for education. An active labor market can, however, also negatively affect school attendance. It is well documented that, as labor market opportunities improve, thus raising the opportunity costs of staying in school, school attendance and performance in school tend to decrease (Neumark and Wascher (1995), Rees and Mocan (1997), Ribar (2001), and Duryea and Arends-Kuenning (2003) for Brazil). It is consequently important to distinguish the contemporary attraction of a favorable labor market from the incentive to pursue schooling when the labor market rewards higher educational attainment (Ribar, 2001).

The third factor that explains educational attainment that should be captured in a structural model is school performance. While school attendance, as a determinant of performance, is an individual and household choice, school performance also depends on factors that are not under the student's and the household's control, such as the child's ability and the quality of the school. This is an important aspect of the educational process since, as Altonji (1993) pointed out, education is a choice made under uncertainty, where ex-ante expectations of academic success influence both enrollment decisions and school attainment. Rochat and Demeulemeester (2001) have provided support to the idea that students consider not only economic returns but also ex-ante chances of success in their educational choices. Estimating a structural model of school attendance and work decisions, Eckstein and Wolpin (1999) were able to identify the role of unobserved ability and motivation from the expected return to education, in the decision for dropping out of high-school. Empirical studies have shown some evidence that school quality affects students' enrollment and achievement (Case and Deaton (1999) for South Africa; Krueger (1999), Card and Krueger (1992, 1996) for the United States; Duflo (2001) for Indonesia; Gould, Lavy, and Paserman (2004) for Ethiopians;

Angrist and Lavy (1999) for Israel), while others have found no relation between school facilities and educational outcomes (Hanushek (1996) and Betts (1995)).

Since the decision process on school attendance and job choice is sequential, with updating and uncertainty revealed at different points in time, its modeling requires a dynamic specification. Comay, Melnik, and Pollatschek (1973) provided one of the first studies that explicitly recognized and formally modeled human capital accumulation as a series of dynamic choices, in which the true benefits of an additional year of schooling include the option value of entry into higher grades. Estimation of structural dynamic models has since been pursued by Cameron and Heckman (1998 and 2001) and by Eckstein and Wolpin (1999). An important feature of the papers by Cameron and Heckman is an explicit control for the dynamic selection of the student body through the grade selection process.

In this paper, we build a structural model of school enrollment and school performance that makes explicit the role of the three factors mentioned above: (1) The "utility" of schooling given by current enrollment, which depends on child and household preferences and current constraints. (2) The expected return to schooling, i.e., the income earning opportunity at each grade level. The return to schooling is itself the combination of a migration decision, a job choice, and the return to schooling in the chosen activity. (3) The expected performance at each grade level.

The choice variable is the annual decision of whether to continue school or not, taken with a dynamic perspective on future choices. School attendance is viewed as a duration process, in the sense that once a child has dropped out of school, he will not return.[1] In that framework, the benefit of going to school is not only the return from one more year of education, but also the possibility of continuing toward higher grades. The

---

[1] There are two reasons why we simplify the model by assuming that children, once they quit school, do not return. The main one is that, in the data set that we have, Progresa has induced an important return to school movement among kids that is only a first year event. As we are interested in a sustainable effect of Progresa, we do not want our estimation to be biased by this first year event. As for the more structural cases of children missing one year of school and returning afterwards, these decisions can only be adjustments for short term fluctuations and shocks that are of a very different nature from the other decisions. Accounting for this requires a very different model (see de Janvry, Finan, Sadoulet, and Vakis, 2004).

expected performance at each grade depends on characteristics of the child and its context.

In our empirical analysis, we assume no unobserved heterogeneity. While this remains a strong assumption with respect to unobserved determinants of schooling choices that may be correlated with the variable of interest, it unfortunately cannot be corrected with cross sectional data. As for the unobserved heterogeneity that only affects the selection process, we argue that it does not bias the calculation of the children's future benefits, nor the decision to enter into secondary school. These results, which are the main purpose of our analysis, are thus unbiased.

3. A structural model of the role of utility, performance, and economic rewards in schooling decisions

In this section, we present a structural model of school and work decisions. The child is assumed to maximize the present value of lifetime utility by choosing at each period whether to continue school or to leave school and go on the labor market. The decision process lasts for a finite number of years until the child reaches his maximum school level. After that, job choice remains the only option. The process is complicated by uncertainty in school performance, which distinguishes the decision to enter a grade from the outcome of successfully finishing the year.

The time line for the decisions and events is illustrated in Figure 1. Consider a child that just graduated from a grade g lower than the maximum grade level offered in school. He can choose to either quit school and go on the job market with a personal qualification given by his school attainment g, or enroll in grade $g + 1$. If he successfully finishes the year in $g + 1$, he faces the next round of choice with completed grade $g + 1$. If he fails, he can either quit school or repeat the grade. After a second failure, however, we assume that the child will not try again.[2]

The main economic reward from schooling is the opportunities that it creates on the job market. Consequently, we assume that the type of job and the possible wage obtained in that job are both functions of school attainment. Let $\pi_j\left(x^e, g\right)$ and $w_j^*\left(x^w, g\right)$

[2] This assumption is consistent with the data that show almost no cases of more than one repetition.

be the probability of obtaining a job of type j and the conditional discounted value of lifetime earnings in job j, respectively, as functions of individual and regional characteristics x^e and x^w, and completed grade g. The expected lifetime income upon quitting school with completed grade g is thus:

$$W_g^*(x) = \sum_j \pi_j\left(x^e, g\right) w_j^*\left(x^w, g\right). \tag{1}$$

The current net utility of being enrolled in grade g, $U_g\left(x^u\right)$, includes the direct utility u of attending school and the transfer T received, net of cost c of schooling, all potentially function of individual (child and household) and regional (village and school) characteristics x^u:

$$U_g\left(x^u\right) = u_g\left(x^u\right) + T_g\left(x^u\right) - c_g\left(x^u\right).$$

Finally, the probability $P_g\left(x^p\right)$ of successfully completing grade g is assumed to depend on personal characteristics x^p of the child and the household.

Two decisions determine the dynamics of school attendance: the decision to repeat in case of failure and the decision to continue school following a success. Let us look first at the decision to repeat. Denote by $EV_{g+1}(x)$ the expected value of lifetime utility of an individual with completed grade $g + 1$, where x includes the elements of x^u, x^p, x^e, and x^w. Consider the general case where the expected direct utility of being enrolled while repeating, U_{g+1}^R, is distinct from the expected utility of first time enrollment, U_{g+1}. The decision to repeat a failed grade compares the expected return of repeating,

$$U_{g+1}^R + P_{g+1} EV_{g+1} + \left(1 - P_{g+1}\right) W_g^*,$$

with the expected income W_g^* that can be obtained from going on the labor market with grade g. The decision to repeat is taken after the child learns of his failure and depends on the child's best forecast of his future benefits. The decision to repeat can be expressed as:

$$R_{g+1} = \Pr\left(\text{repeat } g{+}1\right) = \Pr\left(U^R_{g+1} + \underbrace{P_{g+1}EV_{g+1}}_{\substack{\text{Future Benefits} \\ \text{if success} = FBS_{g+1}}} + \underbrace{\left(1-P_{g+1}\right)W^*_g}_{\substack{\text{Future Benefits} \\ \text{if failure} = FBF_{g+1}}} - W^*_g \geq \varepsilon_g \right). \quad (2)$$

where ε_g represent the unobserved elements of utility or expected income that enter into the child decision. Analogously, the decision to continue school in grade $g + 1$ following a success compares the expected utilities under the decision to continue with the decision to stop school. Denote by $E\left(V_g \middle| s_{g+1}\right)$ the expected utility conditional on the enrollment decision $\left(s_{g+1} \in \{0,1\}\right)$. Collecting all future options gives the expected life time utility of enrolling in grade $g + 1$ as:

$$E(V_g \mid s_{g+1} = 1) = U_{g+1} + \underbrace{P_{g+1}EV_{g+1} + (1-P_{g+1})}$$
$$\underbrace{\left\{R_{g+1}E\left[U^R_{g+1} + P_{g+1}EV_{g+1} + (1-P_{g+1})W^*_g \middle| U^R_{g+1} + P_{g+1}EV_{g+1} + (1-P_{g+1})W^*_g > W^*_g\right] + (1-R_{g+1})W^*_g\right\}}_{\text{Future Benefits} = FB_{g+1}}.$$

The expected future benefits include three terms. The first term, $P_{g+1}EV_{g+1}$, is the probability of success multiplied by the expected lifetime utility with an additional grade. The second term,

$$(1-P_{g+1})\left\{R_{g+1}E\left[U^R_{g+1} + P_{g+1}EV_{g+1} + (1-P_{g+1})W^*_g \middle| U^R_{g+1} + P_{g+1}EV_{g+1} + (1-P_{g+1})W^*_g > W^*_g\right]\right\}$$

is the expected return from failing and repeating. The third term, $\left(1-P_{g+1}\right)\left(1-R_{g+1}\right)W^*_g$, is the expected return from failing and not repeating. The expected utility of not enrolling is:

$$E\left(V_g \middle| s_{g+1} = 0\right) = W^*_g.$$

Similar to the decision on whether to repeat a grade, all decisions are made based on the decision maker's best prediction of his expected future benefits. The decision to continue in school is thus written:

$$S_{g+1} = \Pr\left(s_{g+1} = 1\right) = \Pr\left[U_{g+1} + FB_{g+1} - W^*_g \geq \mu_g\right]. \quad (3)$$

where μ_g represent this unobserved element in this decision.

Our theoretical framework assumes that the unobserved errors $(\varepsilon_g, \mu_g), g = 1, \ldots, G$, are normally distributed with a full covariance matrix. However, for the empirical analysis that we perform, we use a cross-section of observations, with different samples of students taking the decisions to enroll or repeat at each grade level. In this context, with the observations being independent, we cannot estimate the matrix of correlation of errors across grades. The validity of this estimation thus relies on assuming that there is no unobserved heterogeneity in the population of students.

Unobserved heterogeneity can create two important types of biases. One is the standard omitted variable problem, whereby we do not observe some important determinants of school enrollment that are correlated with observed characteristics considered in the model. With cross-sectional data, there is nothing that can be done to correct for this problem, and we have to ignore it as do others (Cameron and Heckman, 1998). The second type of bias comes from the correlation that is created by the mere phenomenon of dynamic selection. Assume for example that school enrollment is related to unobserved ability in addition to a set of observed characteristics. Even if ability is initially not correlated to any of these observed determinants in the population at large, a correlation arises as grade levels increase because low-ability children gradually drop out of school. Cameron and Heckman (1998) showed this to be a source of strong attenuation bias on the role of parents' wealth or education in the schooling decision in upper grades. This dynamic selection bias can be corrected with a non-parametric method proposed by Heckman and Singer (1984) and applied by Cameron and Heckman (1998)[3]. In this paper, we do not correct for this second type of bias either. We justify this as follows. First, we note that unobserved heterogeneity does not affect our calculation of a child's future benefits. These expected values represent a child's best prediction of his future benefits, conditional on his current information set. These predictions are therefore largely based on the observable characteristics and decisions of the older cohorts. Not correcting for dynamic selection is thus likely to create bias in the

[3] The same method is applied by Attanasio, Meghir, and Santiago (2002) in a model similar to this model and applied to the same Progresa sample.

estimation of a child's *current* enrollment decision only. This should, however, not affect the estimation of entry in the first year of secondary school, because there is almost no selection during the primary school years. This is supported by the fact that, in the control villages, less than 10% of 16 years old have discontinued school prior to completing primary school. Moreover, there are no statistically significant differences in the average values of observed characteristics between third graders (the grade at which children begin to drop out) and the children graduating from primary school. Hence this unobserved heterogeneity that only affects the selection process does not bias the estimation of the decision to enter into secondary school, which is the main focus of our analysis.

Weighting the conditional utilities by their choice probabilities gives the unconditional expected lifetime utility of an individual with completed grade *g*:

$$EV_g = S_{g+1}E\left(V_g \middle| s_{g+1} = 1\right) + \left(1 - S_{g+1}\right)E\left(V_g \middle| s_{g+1} = 0\right). \quad (4)$$

Equations (3) and (4) jointly define the recursive process that generates the expected lifetime utility of an individual with any completed grade, which is the value function of the dynamic choice model.

The set of behavioral equations to be estimated is the choice to repeat after a failure (equation (2)) and the choice to continue school after a success (equation (3)). The exogenous information required for these choices is the expected income on the job market at various school attainments, $W_g^*(x)$, and the probability of successfully completing any grade, $P_g\left(x^p\right)$. Before proceeding to these estimations in Sections 5 and 6, we turn to a description of the context of our empirical analysis and document the severity of the schooling problem in Mexico's poor rural population.

4. Data base and descriptive statistics from the control villages

The empirical analysis of this paper is based on data collected for the evaluation of Progresa. An important component of Progresa is the educational subsidies that provide cash transfers to mothers, conditional on their children's regular attendance to school. The program provides a given amount for each child attending school in any of the four upper grades of primary school or the first three grades of secondary school (lower-secondary school). The transfers increase with the grade level and are higher for girls than for boys. These cash transfers range from \$24 to \$30 per month for a child in secondary school, which is roughly half of what a child would earn if working full time,[4] with a maximum of \$74 for a family in 1998. Progresa transfers are important, representing on average 22% of beneficiary families' incomes. By 2000, 2.5 million children were receiving grants.[5]

Eligibility for Progresa was established at the household level, following a two-steps targeting procedure (Skoufias, Davis, and Behrman, 1999). First, poor rural communities were selected on the basis of a marginality index established in 1995 on the basis of information from the Population Census. Then, in these selected communities, a household census was run prior to inception of the program to identify which households are poor, thus becoming eligible. A total of 50,000 communities were selected to receive Progresa and, on average, 78% of the households in the selected communities were declared eligible.

Progresa was implemented following an experimental design in a subset of 506 communities located in seven states. It started to operate in 320 of these communities in May 1998, while the remaining 186 that constitute the control group joined the program in the last round of national incorporation in late 1999. Because transfers are generous, almost all eligible families chose to participate (97%). Hence, while we estimate an intention to treat effect in considering household eligibility rather than effective participation in the analysis that follows, the difference with treatment effect cannot be important. All households in both treatment and control communities were surveyed twice a year from 1997 to 2000.

We use in this paper the first two years of evaluation, which include the baseline census in October 1997, and the two years of follow-up surveys in October 1998 and October 1999 over which the experimental design was maintained. We thus have information on enrollment during three consecutive school years 1997-98, 1998-99, and

[4] The average daily wage of 16-18 years old in the sample is 25 pesos in 1997. A full time work of 20 days per month would generate an income of 500 pesos or \$59 per month.

[5] The Program has been renamed "Oportunidades" and extended to semi-urban areas, with a budget of US\$1.8 billion in 2002.

1999-2000, and on performance in school during academic years 1997-98 and 1998-99. For the econometric analysis, we restrict our interest to the decisions to enroll in the three years of secondary school and in the first year of upper-secondary school since these are the most problematic decisions for school attainment on which Progresa can have an impact. In our sample, this concerns 14,356 children graduating from the last year of primary school or any grade of secondary school.

Statistics on school continuation upon successful completion of a grade (Table 1) show that continuation is almost complete except at entry into secondary school (70.1%) and, as expected, at the end of the lower-secondary cycle (44.6%).

Failure to pass the grade is most severe in the first year of secondary school (Table 2). Failure for the first time occurs in 7-8% of cases in the last year of primary school and in the 2nd and 3rd years of secondary school, with a drop out rate of 15 to 20% among unsuccessful students. Failure in the first year of secondary school is as high as 15%, followed by a drop out rate of 51.9% among those who do not pass the grade. Furthermore, among those that repeat, 24% fail a second time, and for the most part subsequently quit school (68%). Combining these hurdles encountered upon entering into secondary school shows that as many as 36.2% of the students stop school after primary (29.9% never enter secondary school, an additional 5.5% quit after a first failure, and 0.8% after a second failure). This demonstrates that succeeding in secondary school is far more difficult and complex than an aggregate number would suggest. By contrast, failure is low in other grades.

There obviously is heterogeneity in the school population regarding this difficult transition (Table 3). Note in particular the differences between boys and girls (in enrollment, but not in performance, with more boys continuing into secondary), across household heads' education (with more continuation in secondary, less failures, and more repeats in case of failure when the household is educated beyond primary school), and whether a school is available or not in the village. Remarkably, Progresa cancels out the difference between poor and non-poor in continuation in secondary school.

5. Estimation of return to schooling and performance in school equations

5.1. Choice of activity and earnings equations

What is needed for the structural model is a prediction for choice of activity and lifetime earnings that will inform the decision regarding school enrollment. When the time comes for a student to choose a job, he will have additional information on job and wage offers that will inform his choice. Yet, when constructing these anticipated decisions in order to decide whether to continue school, this information is not available. Anticipation of future job choices and conditional earnings in each job are hence based on the outcome of choices made by people that he can observe. The anticipated life-time earnings are, therefore, the weighted average of earnings observed among workers in each job, without correction for any selectivity.

To the extent that occupational choices are made early in one's life, and that they are influenced by the environment at the particular time when this choice is made, we restrict our sample of observations to young adults. This will also better reflect expectations made while still in school. Job choices include activities in the village and migration outside the village for work. We disregard unemployment, since our concern is a long-term choice and not the short-term transition of young people not having found a job yet. Structural underemployment is, however, captured through low monthly earned income. This is particularly important for women, who in many cases combine very low return activities with domestic work. The retained sample for the choice of work in the localities thus includes all the sons and daughters of the household heads, 12 to 18 years old, having left school. For the migration choice, we analyze the decision of all the 13 to 25 years old having migrated for work. Characteristics of the family that are fairly structural such as wealth and parents' education and activities can adequately represent the situation at the time of the decision. Estimation of the migration choice is done separately from that of the occupation choice within the communities for data reasons that we discuss later.

Tables 4a and 4b report a multinomial logit of occupational choice for young men and women that live in the localities. We consider three categories of occupations: agricultural wage earners that represent a large majority of employment (57.7% of all observations), non-agricultural wage earners (34.3%), and self-employed or family labor

(8%). Estimations were done separately for men and women. Indigenous men are substantially less likely to be wage earners than non-indigenous, as do children of households that own land. As expected, education has a strong positive effect on the probability of entering non-agricultural work for both men and women, but is inconsequential for wage work in agriculture. We see the strong influence of parent's own activity, in the sense that, irrespective of gender, children of agricultural workers are more likely to be agricultural workers, and similarly children of non-agricultural workers have a higher tendency to become non-agricultural workers.

For many young people from the poor villages targeted by Progresa, the best job option is to leave their villages and move to more active population centers. Estimation of the migration probability is based on the observed migration between 1997 and 1999, where, for our purpose, migrants are those that left the municipality strictly for employment reasons, and not to study or to get married. As we cannot follow a single cohort from age 13 to 25, we estimated the probability of migrating for three age cohorts, and then compute the overall probability of migration between 13 and 25 years old from these conditional probabilities. Explanatory variables are the same as those used in local activity choice. Note that explanatory variables properly relate to the situation of the individual or his family prior to migration, the time the decision was actually taken. This is made possible by the very large panel data we have in an area of extensive out-migration, which therefore give us enough observations for the estimation of migration behavior. The results on cohort estimations are reported in Appendix Table 1, and a summary of the overall probability of migrating between 13 and 25 years old is reported in Table 5. Results show the important role of education in inducing migration. In addition, boys migrate more than girls. Children from households where the head has lower education, a lower wage, and is an agricultural wage earner migrate more. Children from households with a large number of children and located further away from an urban center also migrate more. These results are all consistent with a Todaro (1969)-type wage gap hypothesis where factors that lower local income opportunities and raise expected income gains at the point of destination increase migration.

To construct life-time earnings, we estimated conditional earnings in each of these occupations by grade level as a function of gender,[6] ethnicity, and age, controlling for distance to an urban center and for state of residence. The estimation is done on the working population 15 years and older. A salient result of these estimations (not reported) is that gender difference is important, with men earning 30% more than women in self-employed jobs[7] and 40% more in non-agricultural wage activities. Indigenous people earn 15% less than non-indigenous in agricultural wage jobs, and 35% less in self-employment, but there is no significant difference in non-agricultural wage jobs. Distance to an urban center is negatively correlated with agricultural wages, but not with earnings in the other two occupations. To predict the level of wages that a village migrant would expect to earn, we resorted to a different source of information, the National Household Income and Expenditure Survey (ENIGH) collected by the National Institute for Statistics, Geography, and Information (INEGI) in 1996. After rejecting the hypothesis that the earning functions are different by grade and by gender, except for intercepts, we performed a single estimation for the whole population. We are unfortunately unable to distinguish between migrants and non-migrants in the ENIGH data, which may be a source of bias. However, since one could argue *a-priori* that migrants do either better or worse than non-migrants, the direction of bias, if any, is unclear. This earning equation is reported in Table 6. The sample consists of individuals 18 and older. As with the other wage equations, the gender difference is very significant, as males earn 17% more than females. The return to education is also very pronounced. The difference in the returns to education between the highest (beyond three years of secondary) and the lowest (no schooling) levels of educational attainment is around 160%.

From these estimations, conditional life-cycle earnings $w_j^*\left(x^w, g\right)$ were predicted for each individual of the school population with characteristics x^w, for each job j, and each potential grade level g, with a discount rate of 5%. Figure 2 shows the sample average of these life-cycle earnings in the four activities. Note that earnings in all three

[6] There was no significant difference between genders except for an intercept. These wage equations are estimated without correction for the selectivity bias of job choices on purpose, because they will be used to predict the expected wage for a young person that would choose the defined job, not for any random person.

local jobs are not only substantially lower than the migration earnings, but also unresponsive to educational level. Education thus has low returns in any occupation in the poorest rural communities targeted by Progresa, while the true reward to education for the youth of these communities is realized through migration, particularly for those who can go beyond three years of secondary school. These results are an eloquent confirmation that educating the youth in poor rural communities is an invitation to migration.

Combining these earnings with the job choice equations gives predicted life-time earnings for each individual with characteristics x and attained grade level g as follows:

$$W_g^*(x) = \pi_{mig}\left(x^e, g\right) w_{mig}^*\left(x^w, g\right) + \left(1 - \pi_{mig}\left(x^e, g\right)\right) \sum_{j \in \text{local}} \pi_j\left(x^e, g\right) w_j^*\left(x^w, g\right),$$

which is a rewriting of equation (1) taking into account the separate estimation of migration and local job choices. These will be used by each individual to compute the return to schooling and hence in deciding on whether to enroll or not in school at different grade levels beyond primary school.

5.2. Performance in school

Table 7 reports the estimation for the probability that a child fails a grade in secondary school, corrected for the fact that we observe failure only among those children that decide to enroll.[8] The relationship between relative age in class (age – (grade + 6)) and probability of failure is complex. On the one hand, being older in class may be a signal of past difficulties and hence lower ability. On the other hand, the variable has its own dynamic: for any given child, the difference can only increase with grade but, across children, the dynamic selection of school attendance probably means that older children drop out of school earlier. In the reported results, the probability of failure decreases with relative age beyond 5th grade. It is interesting to note that, once controlling for individual characteristics and for the decision to repeat a class first failed

---

[7] The self-employed consist of mainly women who supplement their domestic work with low-productive activities.

[8] We also considered separate estimations for whether it was a first time failure or a second time fail. The estimations were much more difficult due to small sample size and resulted in virtually no difference. The reported joint estimation for all failures does control for the impact of repetition on outcomes. Although we correct for selection bias, we eventually predict the conditional probability of failing each grade. Again, we are interested in the probability of failing given that the child enrolls into secondary school and not for a random person.

(selection equation), the probability of failure does not depend on whether it is the first or second time failing that grade. Also interesting is that the performance of poor children (household below poverty line, dwelling with no bathroom) is no worse than that of richer children, and the distance to secondary school, which has a determining impact on enrollment, does not affect performance. Note that Progresa does not influence the probability of failure. This measures the overall effect of Progresa, including the ability of the children that are brought to school by the program and the attendance conditionality associated with the transfers. Children's performance is strongly affected by the presence of someone in the household with a higher educational level. For each additional year of education in the household, the probability of failing decreases by 0.4%. Nine years of education (from no education to a Secondary 3 level) overall decrease the average failure rate by 4.1 percentage points, from 16.1% to 12%.

6. Joint estimation of the enrollment decisions

6.1. Econometric specification

With predicted values for expected lifetime earnings for all school children at every potential grade, and predicted risk of failure at each grade, we can now estimate the dynamic decision to enroll at each grade and, in case of failure, to repeat the grade. There are eight decisions to estimate: to enroll for the first time and eventually to repeat, in Secondary 1 to 3 and in upper-secondary school. This is done successively with backward recursion as follows:

At the terminal level, $G = 10$ (six years of primary, three years of secondary, and one year of upper-secondary),

$$EV_G = W_G^* .$$

For any completed level of education $g < G$, decisions are the following:

a) Consider first whether to repeat $g + 1$ after a first failure:

$$R_{g+1} = \Pr\left(x^u \alpha_{g+1}^R + \beta_{g+1}^R T_{g+1} + \gamma_{g+1}^R FBS_{g+1} + \delta_{g+1}^R FBF_{g+1} + \lambda_{g+1}^R W_g^* \geq \varepsilon_g\right). \quad (2')$$

where:

$FBS_{g+1} = P_{g+1} EV_{g+1}$ is the future benefits in case of success,

and $FBF_{g+1} = \left(1 - P_{g+1}\right) W_g^*$ the future benefits in case of failure.

From this estimation, one derives a future value of repetition conditional on repeating, expressed in units of life-cycle earnings W_g^*:

$$FBR_{g+1} = \left(x^u \alpha_{g+1}^R + \beta_{g+1}^R T_{g+1} + \gamma_{g+1}^R FBS_{g+1} + \delta_{g+1}^R FBF_{g+1} \right) \Big/ \left(-\lambda_{g+1}^R \right). \qquad (5)$$

b) Then consider whether to enroll in the first place after successful completion of the previous grade:

$$S_{g+1} = \Pr\left(x^u \alpha_{g+1} + \beta_{g+1} T_{g+1} + \gamma_{g+1} FB_{g+1} + \lambda_{g+1} W_g^* \geq \mu_g \right) \qquad (3')$$

where:

$$FB_{g+1} = P_{g+1} EV_{g+1} + \left(1 - P_{g+1}\right) \left\{ R_{g+1} FBR_{g+1} + \left(1 - R_{g+1}\right) W_g^* \right\}, \qquad (6)$$

from which one derives the lifetime utility with completed grade *g* in units of life-cycle earnings:

$$EV_g = \left(1 - S_{g+1}\right) W_g^* + S_{g+1} \left(x^u \alpha_{g+1} + \beta_{g+1} T_{g+1} + \gamma_{g+1} FB_{g+1} \right) \Big/ \left(-\lambda_{g+1} \right). \qquad (4')$$

As the parameters $\alpha_g, \beta_g, \alpha_g^R$, and β_g^R of the utility function are grade specific, the model can be estimated recursively. Starting from the highest grade, one knows the expected lifetime utility EV_G with completed grade $g + 1 = G$. First, estimate the repetition probability (2'), in which FBS_{g+1}, FBF_{g+1}, and W_g^* are known predicted values, and the parameters $\alpha_{g+1}^R ,\beta_{g+1}^R, \gamma_{g+1}^R, \delta_{g+1}^R, \lambda_{g+1}^R$ are estimated. Then, use predicted values of repetition to compute FBR_{g+1} (5) and FB_{g+1} (6), and estimate the enrollment decision (3'). Using estimates of parameters $\alpha_{g+1}, \beta_{g+1}, \gamma_{g+1}, \lambda_{g+1}$, and the predicted probability of enrollment S_{g+1}, compute the expected lifetime utility with completed grade *g* with equation (4'). Then, proceed with the next lower grade level.

Note that, in the repetition decision (2'), one can identify separately the three elements of future earnings, namely the expected lifetime earnings if one does not repeat W_g^*, future benefits if one repeats with success, and future benefits if one repeats but fail. This is because future benefits include determinants of the probability of failure that do not enter the earning functions, and future benefits with success include in addition determinants of utility of being in school in future years. Hence, crucial for identification of the estimation is the presence of covariates that explain performance in school and not job choice or earnings, and of covariates that influence preference for school and not probability of failure or earnings.[9] By contrast, in the decision to enroll for the first time, future benefits if successful and upon failure cannot be separately identified, because the possibility to repeat puts all covariates of preferences in the benefits upon failure.

6.2. Estimation results

Table 8 presents partial estimation results, for the enrollment in the first year of secondary school and in upper-secondary school. Covariates that express preferences for school include individual characteristics (gender and rank among children), family characteristics (education, demographics, and indicators of the household welfare level), village characteristics (distance to urban center) and characteristics of the supply of school (distance to secondary school and whether it is a tele-secondary school). Since the choice of which school a child attends is endogenous, the supply side of schooling facility is characterized by the secondary school closest to the village of residence.[10] Although there was an experimental design in the allocation of Progresa transfers, with random selection of control and treatment villages, the effect of Progresa transfers is identified after control for potential bias in the selection process with a Progresa village covariate. While life-cycle earnings were computed in local currency, we use here their logarithms as covariates, and hence future benefits have the dimension of the log of a lifetime income. As mentioned in section 2 above, while our estimation may suffer from a bias due to dynamic selection in the higher grade, the entry into secondary school is not subject to such bias.

Given the small number of children that actually fail, estimation of the decision to repeat is empirically difficult. This problem is further compounded by the fact that for those that fail Secondary 2, almost 90% of them decide to repeat which provides very little variation in the sample (and similarly 84% for Secondary 3). Despite these data limitations, some insights can be gained from the estimation results. Progresa has a strong

---

[9] The child's age is the identifying variable for performance in school, and both the household head's work activity and land holdings are identifying variables for earnings.

[10] A similar approach is taken by Coady and Parker (2003) in their characterization of the supply side of education in the Progresa villages.

positive influence for repeating the first year of secondary school, increasing repetition by 26%. Since the projected life-cycle earnings at the current level of schooling measure the opportunity cost of repeating, it negatively affects enrollment at least in the first year of secondary school. By contrast, the future benefit of an additional year of schooling has a positive impact on the decision to repeat. Also notable is that the poor (household below poverty line, house with no running water) can less afford to have their children repeat grades, while children from larger families can more easily repeat in case of failure.

For first-time enrollment decisions, estimations fare considerably better. As in the grade repetition estimations, poverty (household below poverty line, house with no electricity) reduces enrollment in secondary school. Family characteristics such as maximum education in the household (+), age of the mother (-), and household size (-) are also important determinants. The gender of the child is strongly correlated with the decision to enroll, as boys are more likely to enroll than girls. Additionally, children with higher rank in the family are more likely to enroll. As expected, the distance to secondary school is an important cost and it negatively affects the decision to enroll. The development of a television support system for a rural school is perceived as a positive improvement that induces a higher enrollment rate. Progresa has a positive impact on enrollment in secondary school, increasing the probability of enrolling by 9.8 percentage points. As in the decision to repeat, the wage variable has a strong negative impact, while the future benefit of an additional year of school has a positive effect at each grade level.

The strong results on the role of future benefits in influencing the decision to enroll confirm our presumption of the need to model the school choice as a dynamic process.[11] The advantage of estimating a structural model as we have done here is to clearly identify this role from the current benefits and the opportunity cost of school. As explained above in equations (5) and (4'), these future benefits are normalized to a lifetime earnings (in logarithm). Hence, first time enrollment in secondary school would

[11] Although the theoretical model suggests that we estimate the role of aggregate future benefits for the first-time enrollment, the results for the first estimation (the highest grade) turns out to be very volatile and sensitive to the exact choice of correlates included in the equation. By contrast, when we split the benefits in two, the estimation results are stable. We never encounter this problem for the subsequent estimations of lower grades.

decrease by 5.9 percentage points if the wage at current grade would increase by 10% (W_g^* increases by 0.1) but also increase by 5.7 percentage point if future benefits increase by 10%. The decision to repeat is even more sensitive, with a decline by 15 percentage points for a 10% increase in wage at current grade, and an increase by 13 to 16 percentage points for a 10% increase in future benefits.

7. Impact of committed future CCT and of complementary supply-side programs

7.1. Channels of influence and gains from commitment of future CCT

We decompose the effect of Progresa transfers on the enrollment decision into four channels of influence.

The first and most important channel is the direct utility effect on the enrollment decision of the transfer in the year in which it is received. This effect is measured in the upper part of Table 8. The estimated marginal effect of 9.8 percentage points (pps) is calculated at the average value of other covariates over the whole sample, and not only for Progresa beneficiaries. The corresponding average predicted marginal effect for beneficiaries is 5.6 pps. This would increase the enrollment rate for beneficiaries from the base value of 71.9% (Table 9) to 77.5%.

The second channel is the immediate effect of the transfer on school performance, i.e., the reduction in the probability of failing a grade. This effect is estimated in Table 7 where we see a reduction of 3.7 pps in a 12% failure rate, i.e., a very large (although not precisely measured) 31% decline in grade failure due to the strict school assistance conditionality (children are not allowed to miss more than three school days per month) to qualify for the transfer.

The third channel is the influence of future transfers over the decision to repeat a grade in case of failure. This effect is estimated in the lower part of Table 8. It is large as the transfer induces a marginal effect of 26 pps compared to a base repetition rate of 49.5%, i.e., a 52.5% increase in grade repetition, instead of abandoning school, in case of failure to make the grade.

Finally, the fourth channel is the role of expected benefits from future transfers to be received conditional on continued enrollment through secondary school. This effect is measured jointly with the third effect, as the difference between the predicted total enrollment level with Progresa of 79.2% (Table 9) and the enrollment level induced by the current transfer of 77.5%. This indirect effect, due to credible commitment that future transfers will be received conditional on enrollment, thus accounts for 23.3% of the total effect of Progresa (enrollment increasing from 77.5% to 79.2%), while the direct effect of the immediate transfer accounts for 76.7% of the total effect.

Credible commitment of future transfers, as an element of program design, thus adds nearly a third to the current year transfer in inducing greater contemporaneous enrollment. Committed continuity in spite of political cycles is thus an important determinant of efficiency in CCT programs.

7.2. Impact of a complementary supply-side program reducing the failure rate

An important obstacle for children who enter secondary school is failure to pass the grade, leading to a high rate of dropping out. As seen in Table 2, the failure rate is high, reaching 15% in the first year of secondary school, and 52% of those that fail the grade quit school. Estimation of the determinants of failure highlights the key role of presence in the household of a person with a grade higher than the child. Among children that have completed primary school, only 21% have someone with more than primary education, while 6% do not have anybody with any education, and 33% with only primary education. This suggests the potential benefit of a program designed to help reduce the failure rate of children who have no one with higher education at home. In the spirit of Roemer's (1998) affirmative action interventions to equalize chances of success across children, we predicted the failure rate in Secondary 1 of each child as if they all had someone with Secondary 3 level of schooling at home. The gain is clearly largest for children without an educated person at home. For this group, the probability of success would increase from 79% to 84%, half way to the 88% level of children with educated parents. This sole measure would increase their enrollment rate from 61.8% to 65.9% (Table 9). There is no question that these children's enrollment decisions remain impaired by other characteristics, and hence would still be far below the average enrollment rate, but the program would achieve for that group about half of what the Progresa transfer is currently doing for them (a gain of 8.8 pps). This supply-side program has a greater benefit for non-Progresa beneficiaries (1.5 pps) than for beneficiaries (1.2 pps), simply because the enrollment level of Progresa beneficiaries is initially at a much higher level (79.2% instead of 69.7%). Yet, it complements Progresa by adding 16% to the CCT-induced gain.

7.3. Impact of a complementary supply-side program increasing access to or information about jobs

We saw earlier that the return to education is very low in jobs that are available in poor villages. Yet, outside the village, the return to education is high, particularly for those that reach at least one year of upper-secondary school. Are young people aware of these opportunities and, if so, do they respond to them? We estimated the relationship between predicted probability of migration and predicted wage differential, and found that overall the elasticity is null. There is, however, heterogeneity in the population, and in particular the small group of people that have siblings 13-25 years that migrated during the last 2 years have a wage elasticity of migration of 0.3 to 0.8 depending on their parents' education.[12]

The last simulation is meant to represent the impact of a supply-side program designed to inform young people on job opportunities and conditions outside their village and help them access these jobs. Assuming that such a program will raise the level of responsiveness to that of the most responsive group, we predicted for each child and each level of education the migration rate with an elasticity of 0.8 to wage differential. Increasing the anticipated return to education, in turn, induces greater enrollment. This sole program would increase the enrollment rate by 6 percentage points for children from uneducated households and by 3.3 percentage points for children from the most educated households (Table 9). This program would also increase the enrollment rate by 5.5 percentage points for males and by 5.5 percentage points for the poor. These are of an order of magnitude comparable to the Progresa impact on the targeted populations. Here again, gains are greater for children not benefited by Progresa transfers (5.5 pps) compared to beneficiaries (4.9 pps) because the initial enrollment level is much higher among the latter (79.2% vs. 69.7% for non-beneficiaries). The program however

[12] Other studies find migration elasticities of similar order of magnitude: 1.2 for seasonal farm workers in the U.S. by Perloff, Lynch, and Gabbard (1998); and 0.8 for young people of non-metropolitan counties in the U.S., by Mills and Hazarika (2001).

importantly complements Progresa by adding 67% to the CCT-induced gain in enrollment.

These results show that, while focusing on the demand-side of education through CCT can be effective in enhancing the educational achievements of the rural poor, supply-side programs can also be quite effective. Two such programs that we have shown can have large impacts consist in helping students from uneducated households and in increasing information about or access to job opportunities outside the village. In both cases, the programs have large effect alone, but also add importantly to Progresa gains, showing complementarity between CCT and supply-side programs that modify the context where Progresa is implemented.

8. Conclusions

Progresa is a bold and innovative program introduced to remedy the very low levels of educational achievement among the children of the rural poor, transferring cash to mothers in exchange for sending their children to school with strict attendance requirements. Reduced form estimates of impact have shown that CCT do indeed succeed in raising enrollment levels in secondary school. We use instead here a structural model of the decision to enroll as a function of parents' and children's current utility for school, expected future benefits from going to school, and performance in making the grade. Returns to schooling are the outcome of migration decisions, job choice, and wage earned in the job at given levels of school attainment.

The structural model gives a novel quantitative understanding of how Progresa transfers impact on the current decision to enroll by enabling to decompose this effect into four channels of influence: the influence of current transfers on (1) the utility for schooling and (2) the grade passing performance, and the influence of expected future transfers on (3) the decision to repeat a grade in case of failure and (4) the benefits from continued enrollment. This model allows to simulate the impact on school enrollment of alternative program designs and of program implementation in alternative contexts for which no experimentation has been done.

Decomposition in channels of influence reveals the fundamental importance of committed program continuity on current decisions to enroll. On average across beneficiaries, 23% of the gain in current enrollment is due to expected future CCT.

Indeed, an exceptional merit of Progresa has been its unaltered continuity (except for name change to Oportunidades) across presidential regimes, a rare event in Mexican history.

The structural model also allows to explore the value of supply-side programs additional (and complementary) to the demand-side CCT. One type of program would focus on children from households with no educated adult to help them increase their chances of success in the first year of secondary school. Assuming that the program can compensate for the absence of someone with at least 3 years of secondary school in the household, this alone would increase enrollment rates in secondary school by 1.5 pps among non-Progresa beneficiaries, achieving 16% of the current gain from Progresa for that category of children. For Progresa beneficiaries, the supply-side program adds another 1.2 pps to Progresa's 7.3 pps gain in enrollment in first year of secondary school.

The other type of supply-side program would consist in increasing access to and information about job opportunities outside the village for students from families with no migratory experience. We have seen that having a sibling who migrated is important in raising expected returns to schooling, and thus increases enrollment. Assuming that the program would help all children access the same opportunities and be as responsive to wage differentials as those with older migrant siblings, the enrollment rate would increase by 5.5 pps among non-Progresa children. For Progresa beneficiaries, it adds an extra 4.9 pps to the 7.3 pps gain induced by the CCT.

Constructing structural models of educational choice, rigorously estimated on the basis of randomized treatment, thus allows to experiment with program design and with program impact in alternative contexts modified by complementary programs. Given the very high cost of experimentation, and the political pressures against delaying program implementation, these structural model exercises are fundamental in drawing lessons from implementation that can be used to improve program efficiency.

References

Acemoglu, Daron, and J.-S. Pischke. "Changes in the Wage Structure, Family Income, and Children's Education." *European Economic Review*, May 2001, 45(4-6), pp. 890-904

Akerlof, George, and Rachel Kranton. "Identity and Schooling: Some Lessons from the Economics of Education." *Journal of Economic Literature*, December 2002, 40(4): 1167-1201.

Angrist, Joshua D. and Victor Lavy. "Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement." *The Quarterly Journal of Economics*, May 1999, 114 (2), pp. 533-575.

Attanasio, Orazio, Costas Meghir, and Ana Santiago. "Education Choices in Mexico: Using a Structural Model and a Randomized Experiment to Evaluate Progresa." University College London, November 2002.

Altonji, Joseph. "The Demand for and Return to Education When Education Outcomes are Uncertain." *Journal of Labor Economics*, January 1993, 11(1, Part1), pp. 48-83.

Behrman, Jere, Piyali Sengupta, and Petra Todd. "Progressing through Progresa: An Impact Assessment of a School Subsidy Experiment", forthcoming in *Economic Development and Cultural Change*.

Betts, Julian. "Does School Quality Matter? Evidence from the National Longitudinal Survey of Youth." *The Review of Economics and Statistics*, May 1995, 77(2), pp. 231-250.

Cameron, Steven and James Heckman. "Life Cycle Schooling and Dynamic Selection Bias: Models and Evidence for Five Cohort of American Males." *Journal of Political Economy*, April 1998, 106(2), pp. 263-333.

Cameron, Stephen and James Heckman. "The Dynamics of Educational Attainment for Blacks, Hispanics, and White Males." *The Journal of Political Economy*, June 2001, 109(3), pp. 455-499.

Card, David and Alan Krueger. "Does School Quality Matter - Returns to Education and the Characteristics of Public Schools in the United-States." *Journal of Political Economy*, February 1992, 100(1), pp. 1-40.

Card, David and Alan Krueger. "School Resources and Student Outcomes: An Overview of the Literature and New Evidence from North and South Carolina." *Journal of Economic Perspectives*, Fall 1996, 10(4), pp. 31-50.

Case, Anne and Angus Deaton. "School Inputs and Educational Outcomes in South Africa." *The Quarterly Journal of Economics*, August 1999, 114(3), pp. 1047-1084.

Coady, David P., Parker, Susan W. "Cost-effectiveness Analysis of Demand- and Supply-side Education Interventions: the Case of PROGRESA in Mexico." *Review of Development Economics.* Vol. 8, no. 3 (2004): 440-451

Comay, Y., A. Melnik, and M.A. Pollatschek. "The Option Value of Education and the Optimal Path for Investment in Human Capital." *International Economic Review*, June 1973, 14(2), pp. 421-435.

de Janvry, Alain, Frederico Finan, Elisabeth Sadoulet, and Renos Vakis. "Can conditional cash transfers serve as safety nets to keep children at school and out of the labor market?" University of California at Berkeley, November 2004

Duflo, Esther. "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment." *American Economic Review*, September 2001, 91(4), pp. 795-813.

Duryea, S. and Mary Arends-Kuenning. "School Attendance, Child Labor and Local Market Fluctuations in Urban Brazil." *World Development*, 2003, 31(7), pp. 1165-1178.

Eckstein, Zvi and Kenneth Wolpin. "Why Youths Drop Out of High School: The Impact of Preferences, Opportunities, and Abilities." *Econometrica*, November 1999, 67(6), pp. 1295-1339.

Gould, Eric D., Victor Lavy, and M. Daniele Paserman. "Immigrating to Opportunity: Estimating the Effect of School Quality Using a Natural Experiment on Ethiopians in Israel." *Quarterly Journal of Economics*, May 2004, 119(2), pp. 489-526.

Hanushek, Eric. "Interpreting Recent Research on Schooling in Developing Countries." *World Bank Research Observer*, August 1996, 10(2), pp. 227-246.

Haveman, Robert and Barbara Wolfe. "The Determinants of Children's Attainments: A Review of Methods and Findings." *Journal of Economic Literature*, December 1995, 33(4), pp. 1829-1878.

Krueger, Alan B. "Experimental Estimates of Education Production Functions". *Quarterly Journal of Economics*, May 25, 1999, 114 (2), pp. 497-532.

Mills, Bradford and Gautam Hazarika. "The Migration of Young Adults from Non-Metropolitan Counties." *American Journal of Agricultural Economics*, May 2001, 83(2), pp. 329-340.

Neumark, David and William Wascher. "Minimum-Wage Effects on School and Work Transitions of Teenagers." *American Economic Review*, May 1995, 85(2), pp. 244-249.

Parker, Susan, and Emmanuel Skoufias. "Parental Unemployment, Marital Status Changes, and Child Human Capital: Evidence from Urban Mexico". Washington D.C.: International Food Policy Research Institute, 2000.

Perloff, Jeffrey, Lori Lynch, and Susan Gabbard. "Migration and Seasonal Agricultural Workers." *American Journal of Agricultural Economics*, February 1998, 80(1), pp. 154-64.

Rees, Daniel and H. Naci Mocan. "Labor Market Conditions and the High School Dropout Rate: Evidence from New York State." *Economics of Education Review*, 1997, 16(2), pp. 103-109.

Ribar, David. "The Effects of Local Employment Opportunities on Youth's Work and Schooling." *Economics of Education Review*, August 2001, 20(4), pp. 401-413.

Rochat, Denis and Jean-Luc Demeulemeester. "Rational Choice under Unequal Constraints: The Example of Belgian Higher Education." *Economics of Education Review*, February 2001, 20(1), pp. 15-26.

Roemer, John. *Equality of Opportunity*. Cambridge MA: Harvard University Press, 1998.

Schultz, T. Paul. 2004. "School Subsidies for the Poor: Evaluating the Mexican Progresa Poverty Program". *Journal of Development Economics* 74(1): 199-250.

Skoufias, Emmanuel, Benjamin Davis, and Jere Behrman. "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., 1999.

Székely, Miguel. "Trust the People, Test the Program". *IDBAmérica*, December 18, 2004.

Todd, Petra and Kenneth I. Wolpin. "Using Experimental Data to Validate a Dynamic Behavioral Model of Child Schooling and Fertility: Assessing the Impact of a School Subsidy Program in Mexico". September 2003.

Todaro, Michael. "A Model of Labor Migration and Urban Unemployment in Less Developed Countries." *American Economic Review*, March 1969, 59(1), pp. 138-148.

Willis, Robert and Sherwin Rosen. "Education and Self-Selection." *The Journal of Political Economy*, October 1979, 87(5, Part2), pp. S7-S36.

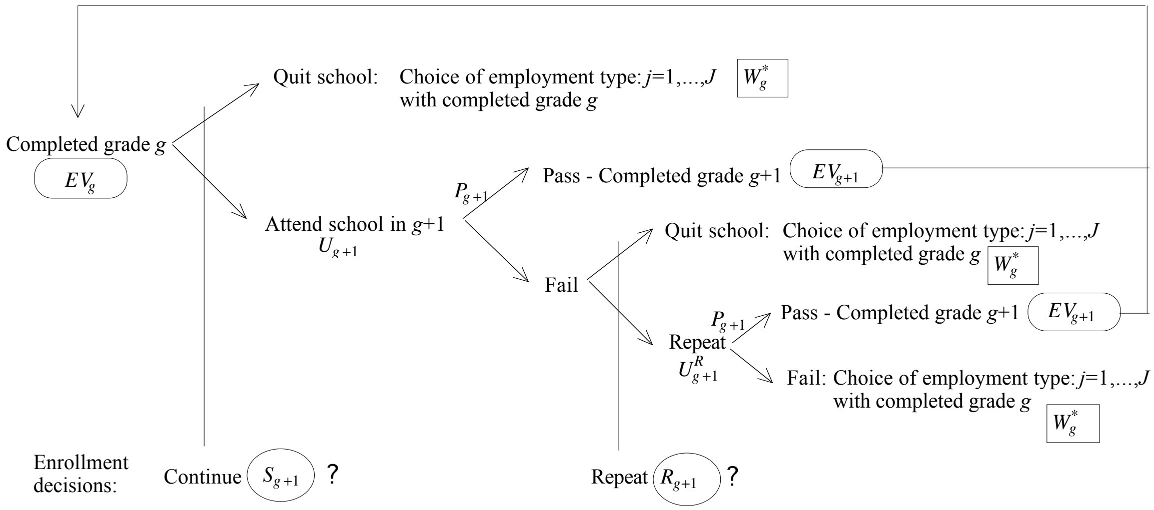

Figure 1. Tree structure of school enrollment choices

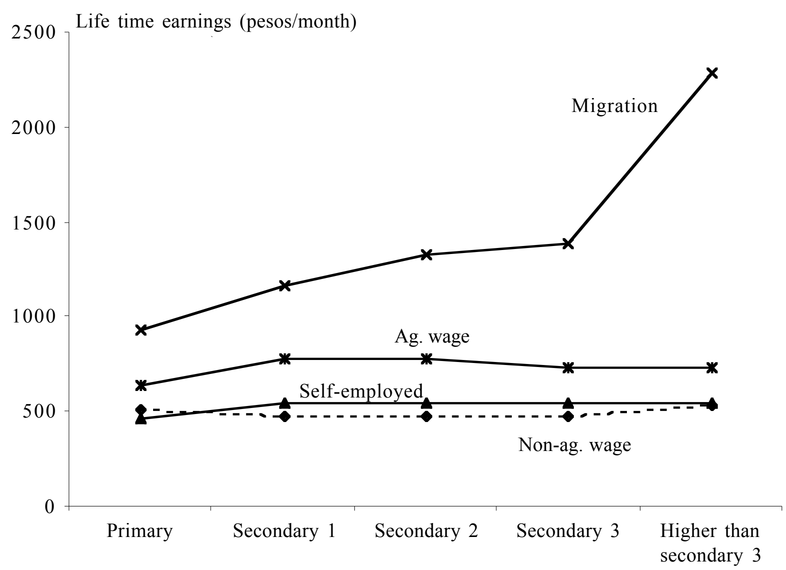

Figure 2. Return to education in the four activities.

Table 1. Schooling decisions after a success

Grade successfully completed in 1997/98	Number of observations	Percentage that enrolled in the next grade in 1998
Primary 4	1,089	96.0
Primary 5	1,173	96.9
Primary 6	1,116	70.1
Secondary 1	571	95.4
Secondary 2	512	96.7
Secondary 3	363	44.6

Observations in control villages.

Table 2. Performance in school and grade repetition

Grade attended in 1998	Grade attended in 1998 for the first time			Grade repeated in 1998		
	Number of observations	Percentage that failed in 1998	Percentage that repeated in 1999 among those that failed in 1998	Number of observations	Percentage that failed in 1998	Percentage that repeated in 1999 among those that failed in 1998
Primary 5	1,045	11.2	95.4	196	13.1	91.3
Primary 6	1,137	7.5	82.3	177	6.6*	72.7*
Secondary 1	782	15.1	48.1	104	23.9	31.8*
Secondary 2	545	7.2	80.6	85	4.0*	100*
Secondary 3	495	8.2	85.3	87	10.1*	75.0*

Observations in control villages.
* Number of observations less than 20.

Table 3. Continuation and performance in secondary school

	Number of observations	Continuation (% of students graduating from primary school)	Failure in Secondary 1 (% of entry)	Repeat after failure in Secondary 1 (% of failing students)	Second failure in Secondary 1 (% of repeat)
All	2,954	75.2	15.5	53.1	18.5
Progresa village	1,838	78.3	15.6	55.8	15.6
Poor	1,180	78.9	15.5	56.5	13.9
Non-poor	658	77.2	15.9	54.4	19.0
Non-Progresa village	1,116	70.1	15.1	48.1	23.9
Poor	702	67.0	15.7	47.0	26.5
Non-poor	414	75.4 **	14.3	50.0	20.9
Girls	1,447	72.8	14.3	54.9	17.2
Boys	1,507	77.4 **	16.6	51.7	19.9
School in village	720	90.0	10.3	60.7	10.2
School not in village	2,234	70.4 **	17.7 **	51.2	21.5 *
Maximum education in household					
None or incomplete primary	493	62.3	24.8	48.5	35.3
Primary	1,700	73.1 **	15.2 **	49.7	19.2 *
More than primary	761	87.5 **	11.8 *	65.7 *	10.1

Observations of continuation in Fall 1998 and performance in school year 1998-99
** (*) significantly different from line above at 1% (5%)

Table 4a. Multinomial estimation of job choice for males 12-18 years old out of school

	Mean value of variable	Agricultural wage earner Relative risk	z of parameter	Non agricultural wage earner Relative risk	z of parameter
Individual characteristics					
Indigenous	0.29	0.53	-3.0	0.56	-2.5
Completed grade					
No schooling	0.31	–	–	–	–
Incomplete primary	0.24	1.03	0.1	0.96	-0.1
Primary	0.45	1.29	0.8	2.09	2.0
Incomplete secondary	0.04	1.64	1.0	2.91	1.9
Secondary or higher	0.20	1.75	1.5	4.85	3.7
Household characteristics					
Father not at home	0.10	2.92	3.5	1.48	1.1
Household head's education	2.03	0.89	-3.0	0.92	-2.0
Household head's wage (pesos per month)	5.62	1.00	-4.4	1.00	-4.5
Household head is ag. wage earner[1]	0.58	19.68	12.6	6.26	7.2
Household head is non ag. wage earner[1]	0.08	11.87	4.1	67.57	6.9
Household head is self employed	0.34	–	–	–	–
Irrigated land per adult (ha)	0.03	2.05	1.0	2.41	1.2
Rainfed land per adult (ha)	0.48	0.90	-2.0	0.78	-3.0
Number of children	3.59	0.99	-0.2	1.07	1.2
Distance to urban center (in km)	102.7	1.01	4.6	1.00	1.6
State controls not reported					
Number of observations	2672	1924		559	
Pseudo R^2	0.24				

[1] Comparison occupation is self-employed.
The relative risk is the exponential of the coefficient. It measures the increase in the relative probability of the category to the base category for a one unit change in the exogenous variable.

Table 4b. Multinomial estimation of job choice for females 12-18 years old out of school

	Mean value of variable	Agricultural wage earner Relative risk	z of parameter	Non agricultural wage earner Relative risk	z of parameter
Individual characteristics					
Indigenous	0.27	1.48	1.1	0.90	-0.3
Completed grade					
No schooling	0.23	–	–	–	–
Incomplete primary	0.16	0.87	-0.3	1.38	0.7
Primary	0.52	0.72	-0.7	2.25	1.8
More than primary	0.25	1.53	0.7	7.10	3.6
Household characteristics					
Father not at home	0.09	0.94	-0.1	0.56	-1.5
Household head's education	2.26	0.88	-2.0	0.94	-1.3
Household head's wage (pesos per month)	5.42	1.00	-0.3	1.00	-0.5
Household head is ag. wage earner[1]	0.43	6.47	5.6	2.36	3.0
Household head is non ag. wage earner[1]	0.12	0.91	-0.2	2.54	2.2
Household head is self employed[1]	0.45	–	–	–	–
Irrigated land per adult (ha)	0.03	0.66	-0.4	1.04	0.1
Rainfed land per adult (ha)	0.34	0.66	-1.6	0.84	-1.0
Number of children	3.79	1.10	1.0	1.27	3.0
Distance to urban center (in km)	96.5	1.01	1.9	1.01	2.8
State control not reported					
Number of observations	948	163		682	
Pseudo R^2	0.22				

[1] Comparison occupation is self-employed.
The relative risk is the exponential of the coefficient. It measures the increase in the relative probability of the category to the base category for a one unit change in the exogenous variable.

Table 5. Estimated probability of migration between 13 and 25 years old as a function of completed grade

	Completed grade					
	No schooling	Primary 6	Secondary 1	Secondary 2	Secondary 3	Post-secondary
Total	14.8	27.2	31.9	29.3	33.6	31.7
By gender						
Boys	17.0	30.6	35.9	32.9	37.4	35.8
Girls	12.4	23.5	27.6	25.4	29.4	27.3
By occupation of head of household						
Ag. Worker	16.2	29.4	34.0	31.6	35.9	33.4
Non-age worker	11.8	22.8	27.5	24.7	28.7	27.7
Self-employed	13.7	25.5	30.3	27.6	31.6	30.3
By head of household education level						
Primary	11.1	21.9	25.8	23.6	27.6	25.6
More than primary	6.7	14.5	17.3	15.7	18.9	17.3

Predicted probability calculated by sample enumeration from the estimation reported in Appendix Table 1.

Table 6. Wage equation for migrants
(Endogenous variable is log of monthy wage)

	Mean value of variable	Coefficient	z
Wage (peso/month)	889		
Individual characteristics			
Gender (male = 1)	0.66	0.167	9.2
Age	37.40	0.069	6.2
Age^2 (/1000)		-8.6E-04	-3.3
Age ^3 (/1000000)		1.6E-06	0.9
Completed grade			
No schooling	0.12	–	–
Incomplete primary	0.24	0.308	9.9
Primary	0.21	0.682	20.8
Secondary 1	0.02	0.912	12.3
Secondary 2	0.03	1.042	17.2
Secondary 3	0.19	1.081	31.3
Higher than secondary 3	0.20	1.584	46.8
Intercept	1	4.486	29.2

Estimated from ENIGH, 1996.
Number of observations: 17,403. $R^2 = 0.19$.

Table 7. Probability of failing grades in secondary and post-secondary school, school year 1998-9

	Mean value variable	Marginal effect (%)	z of coefficient
Failure (%)	12.0		
Individual characteristics			
Age – Grade + 6	-0.353	11.7	3.8
(Age – Grade + 6)*Grade	-2.591	-1.9	-4.7
Repeating	0.120	0.6	0.5
Progresa transfer received (0/1)	0.5	-3.7	-1.3
Attended grade			
Secondary 1	0.400	9.8	3.6
Secondary 2	0.277	-2.0	-1.5
Secondary 3	0.240	–	–
Upper-secondary 1	0.083	24.6	4.2
Household characteristics			
Household maximum education	6.20	-0.4	-2.8
Dwelling has bathroom (0/1)	0.70	-1.0	-1.1
Poor	0.79	1.5	0.9
Progresa village	0.64	2.7	1.9
Distance to secondary school (km)	1.93	0.5	1.2
Intercept	1		-13.3

The selection equation of school attendance includes three identification variables (house has a dirt floor, gender, and family size). None of them are significant when introduced in the failure equation and, jointly, their non-significance cannot be rejected (Chi2(3) = 4.94, p-value=.18).
Number of observations: 6228 attending school. Wald chi2 (10) = 126.1.
Standard errors are adjusted for clustering on households.

Table 8. Enrollment decision in Secondary 1 and Upper-secondary 1, school years 1998-99 and 1999-2000

	Secondary 1			Upper-Secondary 1		
	Mean value variable	Marginal effect	z of coefficient	Mean value variable	Marginal effect	z of coefficient
First-time enrollment decision (equation (3'))						
Observed percentage enrollment	75.9			44.8		
Current utility U(*x* ")						
Gender (male = 1)	0.51	0.060	2.6	0.55	0.177	3.8
Rank among children	2.26	0.024	3.6	1.63	-0.013	-0.7
Household maximum education (years)	5.59	0.017	6.9	6.61	0.021	4.7
Family size	7.31	-0.015	-4.4	7.02	-0.010	-1.5
Mother's age	38.77	-0.003	-3.0	40.63	-0.006	-3.6
Dwelling has dirt floor (0/1)	0.58	0.011	0.9	0.49	-0.017	-0.6
Dwelling has electricity (0/1)	0.76	0.049	3.5	0.85	0.084	2.4
Poor (0/1)	0.85	-0.051	-1.98	0.74	0.031	0.97
Distance to secondary school (km)	2.26	-0.027	-7.1			
Tele-secondary school (0/1)	0.83	0.058	3.6			
Progresa village	0.61	-0.034	-1.1	0.59	0.023	0.9
Progresa transfer *T* (0/1)	0.53	0.098	2.90			
Future benefits (ln of pesos)						
Future benefits/success *FBS*				5.04	0.539	2.0
Future benefits/failure *FBF*				6.76	0.185	2.2
Furture benefits *FB*	7.40	0.565	6.5			
Wage with current schooling W^* (ln of pesos)	6.50	-0.591	-9.4	6.85	-0.973	-3.2
Number of observations		5272			1729	
Pseudo R^2		0.13			0.03	
Decision to repeat, in case of failure (equation (2'))						
Observed percentage enrollment	49.5			43.0		
Current utility U(*x* ")						
Gender	0.50	0.008	0.1	0.54	0.156	1.2
Household maximum education (years)	5.39	-0.006	-0.8	7.03	0.036	3.2
Family size	7.39	0.028	2.8	6.86	-0.018	-1.0
Number of rooms in house	1.95	0.011	0.6	2.41	0.002	0.1
Dwelling has dirt floor (0/1)	0.78	0.008	0.2	0.88	-0.017	-0.2
Dwelling has electricity (0/1)	0.43	0.091	2.2	0.51	-0.064	-0.9
Indigenous (0/1)	0.28	-0.075	-1.4	0.24	0.129	1.6
Poor (0/1)	0.82	-0.178	-2.1	0.65	0.010	0.1
Distance to secondary school (km)	2.35	0.016	1.2			
Distance to an urban center (km)	101.9	-0.001	-1.1	105.1	0.001	0.6
Progresa village (0/1)	0.62	-0.169	-1.7	0.64	0.118	1.7
Progresa transfer *T* (0/1)	0.5	0.260	2.5			
Future benefits (ln of pesos)						
Future benefits/success *FBS*	6.22	1.552	4.1	4.58	0.904	1.2
Future benefits/failure *FBF*	1.17	1.333	3.2	2.48	1.031	1.3
Wage with current schooling W^* (ln of pesos)	6.52	-1.505	-5.7	6.83	-1.288	-1.4
Number of observations		792			272	
R^2		0.14			0.06	

Marginal effects computed at average value of covariates. For dummy variables, marginal effects are for discrete change from 0 to 1. Results for Secondary 2 and Secondary 3 not reported.

Table 9. Effects of complementary programs on enrollment of children having achieved primary school

	Number of observations	Predicted enrollment in secondary		
		Without program (%)	With program (%)	Difference (%)
Performance support program[1]				
All children	5272	76.0	77.2	1.2
By maximum education in household				
No education	332	61.8	65.9	4.1
Primary school	1763	76.2	77.0	0.8
				**
By random assignment				
Progresa beneficiaries	2769	79.2	80.4	1.2
Non-Progresa beneficiaries	1708	69.7	71.2	1.5
				**
Information on job opportunities (increasing the wage elasticity of migration to 0.8)				
All children	5272	76.0	81.3	5.3
By gender				
Males	2583	78.4	83.9	5.5
Females	2689	73.4	78.7	5.2
				**
By maximum education in household				
No education	332	61.8	67.8	6.0
Upper-secondary school	320	89.6	92.9	3.3
				**
By welfare level				
Non-poor	795	77.9	82.8	4.9
Poor	4477	75.6	81.1	5.5
				**
By random assignment				
Progresa beneficiaries	2769	79.2	84.3	4.9
Non-Progresa beneficiaries	1708	69.7	75.9	5.5
				**

[1] Failure rate in first year of secondary school set as if maximum education in household was Secondary 3.
** significantly different at .001

Appendix Table 1. Probit estimation of migration, by cohort

	13–16 years old			17–20 years old			21–25 years old		
	Mean value variable	dF / dx (%)	z	Mean value variable	dF / dx (%)	z	Mean value variable	dF / dx (%)	z
Individual characteristics									
Male	0.52	1.0	2.8	0.51	1.6	3.8	0.49	1.2	4.3
Age in 1997	14.5	1.2	7.9	18.4	0.3	1.7	23.0	-0.1	-1.4
Indigenous	0.29	0.8	1.6	0.28	0.3	0.6	0.29	-0.3	-0.9
Male*indigenous	0.15	-2.0	-3.7	–	–	–	–	–	–
Completed grade									
No education	0.04	–	–	0.08	–	–	0.11	–	–
Incomplete primary	0.24	1.1	1.1	0.21	6.1	3.8	0.27	0.5	0.8
Primary 6	0.36	1.7	1.8	0.38	6.6	4.7	0.39	1.3	2.1
Secondary 1	0.12	1.2	1.1	0.02	11.5	3.6	0.01	3.3	1.8
Secondary 2	0.12	2.4	2.1	0.03	8.5	3.3	0.02	2.4	1.6
Secondary 3	0.10	2.6	2.2	0.20	11.2	6.0	0.16	2.4	3.0
Higher than secondary	0.02	-0.3	-0.2	0.08	12.9	5.6	0.04	3.1	2.7
Household characteristics									
Father not at home	0.16	1.2	2.8	–	–	–	–	–	–
Household head's education	2.47	-0.2	-2.6	2.41	-0.3	-3.4	3.07	-0.3	-4.9
Household head's wage (pesos per month)	575	0.0	-2.3	566	0.0	-1.9	601	0.0	-0.6
Household head is ag. wage earner	0.48	1.0	2.9	0.46	0.4	0.8	0.49	0.0	0.1
Household head is non ag. wage earner	0.09	0.3	0.4	0.08	1.0	1.0	0.09	0.2	0.4
Household head is self employed	0.44	–	–	0.46	–	–	0.42	–	–
Number of children	3.68	0.4	4.6	2.88	0.7	6.1	1.46	0.6	8.1
Irrigated land per adult (ha)	0.04	0.5	1.9	0.02	-0.2	-0.2	0.02	0.3	0.5
Rainfed land per adult (ha)	0.55	0.0	0.4	0.43	-0.3	-1.1	0.37	-0.2	-1.2
Distance to urban center (in km)	103.6	0.0	7.1	100.7	0.0	4.8	100.3	0.0	4.4
State control									
Guerrero	0.08	–	–	0.08	–	–	0.08	–	–
Hidalgo	0.17	6.4	4.7	0.17	6.5	4.0	0.17	3.2	2.8
Michoacan	0.13	8.3	5.2	0.12	9.6	4.9	0.12	7.5	4.6
Puebla	0.16	6.8	4.6	0.16	6.2	3.7	0.16	2.6	2.2
Queretaro	0.06	8.4	4.5	0.07	10.4	4.8	0.06	8.3	4.5
San Luis Potosi	0.16	6.1	4.4	0.16	7.7	4.5	0.16	4.2	3.4
Veracruz	0.25	8.5	6.3	0.24	8.5	5.3	0.24	4.7	4.0
Number of observations	13205			10122			9321		
Observed percentage migrating	4.12			5.83			3.01		
Pseudo R2		0.06			0.05			0.09	