The role of labor market opportunities in explaining school attendance
by

Elisabeth Sadoulet, Frederico Finan, and Alain de Janvry
Department of Agricultural and Resource Economics
University of California at Berkeley

Draft - September 2001

1. What motivates children to attend school

Understanding what motivates children to attend school is essential for the design of policies that attempt to increase the level of education. Many empirical studies have related a child's schooling to his own or his parents' characteristics (Haveman and Wolfe, 1995; Acemoglu and Pischke, 2001). It is possible that the schooling of children provides some current utility to the family, for instance because educated parents prefer educated children. Schooling can thus be viewed as a "consumption good" that is purchased according to preferences and under a budget constraint. Increasing school attendance can thus be obtained by campaigns to alter parents' preferences, by direct transfers to relax their budget constraint, or by conditional transfers that reduce the cost of schooling. Alternatively, schooling can be viewed as an investment for future earnings, as documented in the early work of Willis and Rosen (1979). In this case, attendance should be determined by the return to schooling and, if the capital market works, it is only to the extent that children's and parents' characteristics affect the return to schooling that they should affect attendance to school. As a first approximation, school attendance would mostly be affected 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 schooling is not easily implementable, transfers or conditional transfers would help households close the gap with their optimal investment. Note, however, that it is the perceived rather than the market return to schooling that dictates the schooling decision, and hence imperfect information on or unequal access to job market opportunities may substantially reduce the incentives for a child to get a higher educational level. This suggests that promoting in school access to information about job opportunities may encourage 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, which raises the opportunity costs of staying in school, school attendance and performance in school decrease (Neumark and Wascher (1998), Rees and Mocan (1997), Ribar (2001), Kane (1994), although Card and Lemieux (1997) shows no influence on women schooling). It is important to distinguish the contemporary attraction of a favorable labor market with the incentive to pursue schooling when the labor market rewards higher education attainment (Ribar, 2001).

Another important dimension is to distinguish school attendance from school performance. While school attendance is a household choice, school performance depend on factors not under 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) has pointed out, education is a choice made under uncertainty, where ex-ante expectations of academic success influence both school attainment and occupational choice. Recently, Rochat and Demeulemeester (2001) have provided support to the idea that students consider both economic returns and 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; Card and Krueger (1992, 1996)). However, as Hanushek (1996) and Betts (1995) have pointed out, the evidence is far from conclusive, as both studies show no relation between school facilities and economic outcomes. These results point to the importance of policy on improving school quality and on designing academic support for children with difficulties in order to increase children's performance and enrollment.

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 model. Comay, Melnik, and Pollatschek (1973) provided one of the first studies that explicitly recognized and formally modeled human capital accumulation as a series of complex and dynamic choices. Using a dynamic programming approach, they showed that 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 developed by Heckman (1998) and Cameron (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 students body through the grade selection process.

In this paper, we build a structural model of school attendance and school performance that identifies three factors:

- The "utility" of schooling given by current enrollment, which depends on preferences and current constraints.
- The expected performance at each grade level.
- The return to schooling, i.e., the income earning opportunity at each grade level. The return to schooling is in itself the combined result of an activity choice and the return to schooling in the chosen activity.

The choice variable is the annual decision of whether to continue school or not, taken with a dynamic perspective of 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. In that framework, the benefit of going to school is not only the return of one more year of education, but also the possibility of continuing toward higher grades.

The structural model is then used to identify different policy instruments that could improve school attendance and achievement. We examine in particular the role of Progresa, a program that transfers resources to households for sending their children to school. Progresa payments are expected to influence school attendance through two channels. The first is that the transfer itself attaches some current utility to school attendance. In a dynamic perspective, attending school is induced not only by the current transfer received but also by the perspective of future transfers. The second channel is that, by requiring a steady school attendance (children are not allowed to miss more than three days per month to receive the transfers), Progresa should improve the performance in school. Better performance in turn induces greater attendance.

We identify two other types of instruments: those improving the performance in school, and those improving information on job opportunities and on the return to schooling. Critical to the performance in school is the quality of the school. However, the more resilient factor we find affecting children's performance is their parents' own education. One can, therefore, ask oneself what programs could be used to compensate for the differential handicap that children from noneducated parents face, a question identical to Romer's (2001) concept of affirmative action to equalize chances of success. Another interesting finding of this study is what seems to be a very suboptimal choice of jobs and, as a consequence, a very low return to schooling. One notes in particular a low correlation between job choices and potential lifetime earnings in different jobs. This suggests the existence of imperfect information on jobs that could be obtained with higher educational levels. This perceived low return to schooling in turn detracts children from pursuing education further.

2. Theory: A dynamic model of school and work decisions

In this section, we present a dynamic 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 the maximum school level. After that, job choice remains the only option. The process is complicated by the uncertainty in school performance, which distinguishes the decision to enter a grade and the result, which is not under the child's control, 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 grade g lower than the maximum grade level offered in school. He can choose either to quit school and go on the job market, with a personal qualification given by his school attainment g, or to 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 cannot repeat, as this corresponds to observations made in the field.

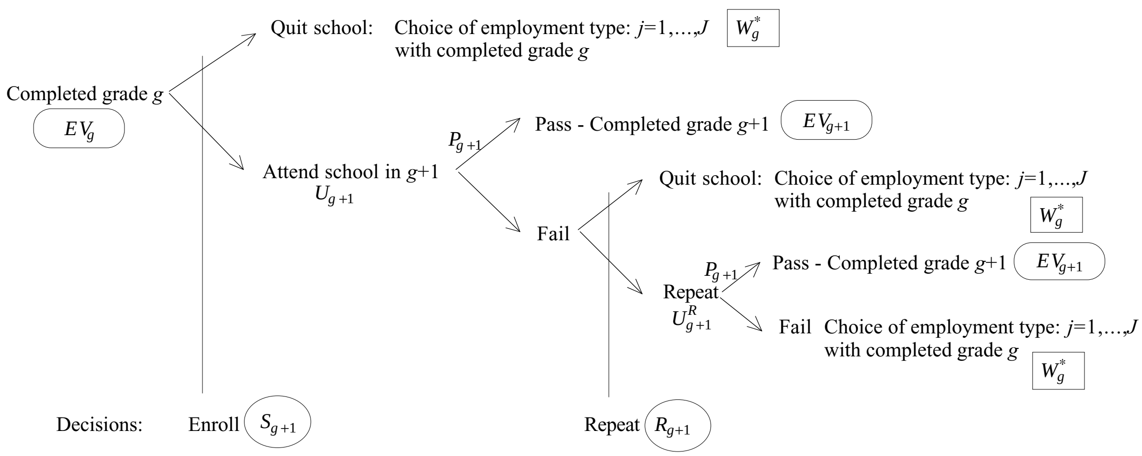

Figure 1. Tree structure of school enrollment choices

The main economic reward from schooling is the opportunities that it opens on the job market. This is captured by assuming that the job that can be obtained as well as the wage in that job category are function of school attainment. Let $\pi_j\left(x^e,g\right)$ and $w_j^*\left(x^w,g\right)$ be the probability of obtaining a job of type j and the discounted value of 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 school in grade g, $U_g\left(x^u\right)$, includes the direct utility u of attending school and transfer T received, net of cost c of schooling, all potentially function of individual and regional 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.

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 $g + 1$. Consider the general case where the direct utility of being enrolled when repeating, U_{g+1}^R, is distinct from the utility at first try U_{g+1}. The decision to repeat a failed grade proceeds from the comparison of the utility obtained by 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. Let ε be an error term in the comparison process, the decision will be the following:

$$R_{g+1} = \Pr(\text{repeat g} + 1) = \Pr\left(U_{g+1}^{R} + \underbrace{P_{g+1}EV_{g+1}}_{\substack{\text{Future Benefits} \\ \text{if success } FB/S_{g+1}}} + \underbrace{\left(1 - P_{g+1}\right)W_g^*}_{\substack{\text{Future Benefits} \\ \text{if failure } FB/F_{g+1}}} - W_g^* \geq \varepsilon \right). \tag{2}$$

The decision to continue school in grade g + 1 following a success will proceed from the comparison between the expected utilities under the decison to continue and to stop school. Denote $E\left(V_g \middle| s_{g+1}\right)$ the expected utility conditional on the enrolment decision $\left(s_{g+1} = 0/1\right)$. Collecting all future options gives the expected life time utility of enrolling in grade g + 1 as:

$$E\left(V_g \middle| s_{g+1} = 1\right) = U_{g+1} + \underbrace{P_{g+1}EV_{g+1} + \left(1 - P_{g+1}\right)\left\{R_{g+1}\left(U_{g+1}^{R} + P_{g+1}EV_{g+1} + \left(1 - P_{g+1}\right)W_g^*\right) + \left(1 - R_{g+1}\right)W_g^*\right\}}_{\text{Future Benefits} = FB_{g+1}},$$

while the expected utility of not enrolling is:

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

If μ is the error term in the comparison process, the decision to enroll in school is:

$$S_{g+1} = \Pr\left(s_{g+1} = 1\right) = \Pr\left[U_{g+1} + FB_{g+1} - W_g^* \geq \mu\right] \tag{3}$$

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

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

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 are the expected income on the job market with various school attainments $W_g^*(x)$ and the probability of failing any grade $P_g\left(x^p\right)$.

3. The empirical strategy

1. *Choices of job type and earnings* are estimated for given school attainment and exogenous personal and regional characteristics. Covariates x^w for the wage equation will typically include gender and regional

characteristics of the job market. Covariates x^e for the job choice equation will in addition include parent's activities.

Note that we need the probability of entering in a given job type for given educational level and family or environmental characteristics. Estimating job choices on the basis of the adult population in Progresa villages is problematic for two reasons. One is that we certainly have a selected population that stayed in the villages, while better opportunities are available outside, in non-marginal areas or in cities. Second, we do not observe the characteristics of the individuals at the time they made their choice, but now at the time of the survey. Notably, one does not know the wealth or the activity of the parents of those that are currently head of households. To mitigate these two problems, we include "leaving the village" among job choices, and we only consider young people that have recently left school. The sample for this choice is therefore all the sons and daughters of the head of the households, 13 to 25 years old, having left school, and either living in the household or having migrated. Characteristics of the family that are fairly structural such as wealth and parents' education and activities can be considered as adequately representing the situation at the time of the decision.

For the wage equation, estimation of profiles of local earnings as a function of activity, grade, and age allows for the computation of a discounted life-time earning. In order to capture the level of wages that a village migrant would expect to earn, we resort to another data set, the National Household Income and Expenditure Survey (ENIGH) collected by the Mexican Statistical Institute INEGHI in 1996. With the different data set we were unable to make the wage equations entirely comparable, as a result these outside the village earnings are estimated only as a function of gender, grade and age. We were also unable to distinguish between immigrants and non-immigrants in the ENIGH data set, which may be a source of upward bias.

2. *Failure* is exogenous to the schooling choices in this model. The probability of failure can be estimated with our data, as a function of individual and household characteristics, and school characteristics. These school characteristics are obtained from a separate school survey, and to avoid any endogeneity problem of school choice, each child has been assigned the school closest to its community. Failure is particularly important for the first year of secondary school. While our model allows for a child to fail twice, we do not distinguish between a first failure and a second failure. We found that little is gained by making this distinction.

3. The estimation of the two decisions choices can then be done as follows.
At the terminal level, G = 10,

$$EV_G = W_G^*.$$

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

Repeat $g + 1$ after a first failure:

$$R_{g+1} = \Pr\left(x^u \alpha_{g+1}^R + \beta_{g+1}^R T_{g+1} + \gamma_{g+1}^R FB/S_{g+1} + \delta_{g+1}^R FB/F_{g+1} + \lambda_{g+1}^R W_g^* \geq \varepsilon\right). \tag{2'}$$

where:

$$FB/S_{g+1} = P_{g+1} EV_{g+1}\ ,$$
$$FB/F_{g+1} = \left(1 - P_{g+1}\right) W_g^*,$$

Enroll in the first place:

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

where:

$$FB_{g+1} = P_{g+1} EV_{g+1} + \left(1 - P_{g+1}\right)$$
$$\left\{R_{g+1}\left(x^u \alpha_{g+1}^R + \beta_{g+1}^R T_{g+1} + \gamma_{g+1}^R FB/S_{g+1} + \delta_{g+1}^R FB/F_{g+1}\right) + \left(1 - R_{g+1}\right) W_g^*\right\}. \tag{5}$$

From which one derives the life time utility with completed grade g:

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

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 upper grade, one knows the expected lifetime utility EV_G with completed grade $g + 1 = G$. First estimate the repetition probability (2'), in which FB / S_{g+1}, FB / F_{g+1}, and W_g^* are known predicted value, 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 FB_{g+1} (5), 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.

4. Data base and descriptive Statistics from the control villages

The data that we use in this paper come from a data collection effort undertaken for the evaluation of a large welfare program, the Education, Health, and Nutrition Program, known by its Spanish acronym Progresa, in rural Mexico. The purpose of the program is to provide resources and incentives to increase the human capital of the children of poor rural households, thus attempting to break the inheritance of poverty. The program provides cash transfers to poor households, conditional on the child's school attendance and on regular visit to health centers. Overall, these cash transfers represent 22% of the income of the beneficiary families. The program has grown rapidly since its inception in 1997 and now covers 2.6 million rural families in extreme poverty, corresponding to about 40 percent of all rural families in Mexico. Progresa currently operates in 50,000 localities in 31 states, with a budget of approximately one billion dollars for 2000.

Eligibility for the program is established at the household level. The Progresa program operates in all poor communities (defined by a national marginality index developed from the 1995 census) that have minimal access to primary school and primary care facilities, and all households characterized as poor in these communities are eligible (see Skoufias, Davis, and Behrman, 1999). Households' poverty status has been established prior to the start of the program from an exhaustive household survey run in the poor communities in 1997 (ENCASEH 1997 survey). A total of 50,000 communities were targeted to receive Progresa and, on average, 78% of the population of the selected communities is eligible.

For purposes of evaluation, the Progresa program has been implemented following an experimental design. A subset of 506 communities were selected to participate in the evaluation. Each of these communities was randomly assigned either to the treatment group where Progresa was implemented starting in 1998, or to the control group where Progresa would be introduced three years later (Behrman and Todd, 1999). All households (eligible and non-eligible) of both types of communities were then surveyed twice a year during the three years of the evaluation. These experimental communities are located in seven states (Guerrero, Hidalgo, Michoacan, Puebla, Queretero, San Luis Potosi, and Veracruz). There are 320 treatment localities and 186 control localities in the experiment. Program benefits began in May 1998. The unbalanced sample including all individuals present at some point in time between October 1997 and November 1999 is of 152,000 individuals from 26,000 households. Because transfers are generous, almost all eligible families chose to participate (97%).

We use in this paper the first two years of evaluation, which include the base line survey (ENCASEH 97), and the follow up surveys in October 1998 and October 1999. We thus have information on enrollment during three consecutive school years 1997-98, 1998-99, and 1999-2000, and on performance in school during the academic years 1997-98 and 1998-99. Our total sample includes xxxx children at different grade levels of primary and junior high school.

Statistics on school continuation upon successful completion of a grade (Table 1): Shows that continuation is almost complete except at the entry of secondary school and, as expected, at the end of the junior high cycle.

Table 1. Schooling decision after a success

Successfully completed grade in 1997/98	Observations	Enrolled in 98 (%)	Stopped (%)
Primary 4	1,089	96.0	4.0
Primary 5	1,173	96.9	3.1
Primary 6	1,116	70.1	29.9
Secondary 1	571	95.4	4.6
Secondary 2	512	96.7	3.3
Secondary 3	363	44.6	55.4

Observations in control villages

Failures and repetition (Table 2). Failure is again a more severe problem in the first year of secondary school. Failure for the first time occurs in 7-8% of the case in the last year of primary school and the 2nd and 3rd year of secondary school, with a drop out rate of 15 to 20% among the unsuccessful students. Failure in the first year of secondary school is as high as 15% and followed by a drop out rate of 51.9%. Furthermore among those that repeat their year, 24% fail a second time, and for the most part quit school after. Combining these hurdles encountered upon entering into secondary school, 36.2% of the students stop school after primary school (29.9% never enter secondary school, an additional 5.5% quit after a first failure, and 0.8% after a second failure). This shows the process of entering secondary school as far more difficult and complex than an aggregate number would suggest. Second failure almost never occurs in the other grades.

Table 2. Performance in school and grade repetition

	Grade attended in 1998 for the first time				Grade repeated in 1998			
			Among those that failed in 98				Among those that failed in 98	
Grade attended in 1998	Observations	Failed in 98 (%)	Repeat in 99 (%)	Quit (%)	Observations	Failed in 98 (%)	Repeat in 99 (%)	Quit in 99 (%)
Primary 5	1,045	11.2	95.4	4.6*	196	13.1	91.3	8.7*
Primary 6	1,137	7.5	82.3	17.7*	177	6.6*	72.7*	27.3*
Secondary 1	782	15.1	48.1	51.9	104	23.9	31.8*	68.2
Secondary 2	545	7.2	80.6	19.4*	85	4.0*	100*	0
Secondary 3	495	8.2	85.3	14.7*	87	10.1*	75.0*	25.0*

Observations in control villages
* Number of observations less than 20

Contrast across groups (Table 3): Note difference between boys and girls (in enrollment, but not in performance), across head of household's education. Most interestingly, Progresa cancels out the difference between poor and non-poor.

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					
School not in village					
Education of the head of household					
None or incomplete primary	1,481	69.0	18.5	48.5	28.4
Primary	1,335	80.0	13.1	55.1	12.8
More than primary	138	94.2	11.7	92.3	0.0
Family size					
Three or less children 0-14 years old	1,441	75.3	15.4	53.4	17.7
More than three children 0-14 years old	1,513	75.1	15.5	52.8	19.1
Birth order					
Child is oldest in family	976	73.7	18.0	50.0	18.3
Child not oldest	1,978	75.9	14.3	54.9	18.6

Observations of continuation in Fall 1998 and performance in school year 1998-99

5. Econometric estimation

We assume:

- no re-entry possible after school is dropped
- no third repetition of a class (as in the model)

5.1. Choice of activity and earnings equations

What is needed here is a prediction of choice of activity and life time earnings that will inform the decision regarding optimal school achievement. To the extent that occupational choice are made early in the life of an individual, and that it is influenced by the environment of the individual at the time of his choice, we restrict our sample to the young adults. This will also better reflect expectations made while still in school. Ideally we would like to have all young people. However, given the importance of the family environment and in particular parents' own activity on the choice of activity, we have to restrict further our sample to the young adult that are still in their parents' household (since we do not have family background information on those that are head of households). Table 4 reports a multinomial logit of occupational choice.

We only consider three occupations for those who stay in the localities, agricultural wage earners who represent a large majority of employment with 57.7% of all observations, non-agricultural wage earners, and self-employed or family labor. Men are more likely to have an agricultural employment and less likely to be non-agricultural workers than women. Indigenous people are less likely to be agricultural wage earners than non-indigenous. As expected, education has a strong positive effect on the probability of entering non-agricultural work. Children of household that own land are, as expected, less likely to be wage earner. We see the strong influence of parent's own activity, in the sense that kids of agricultural workers are more likely to be agricultural workers than the other, and similarly kids of non-agricultural workers have higher tendency to become non-agricultural workers. Distance to urban center has the unexpected positive impact on the probability to have agricultural wage work or be a non-agricultural wage earner.

Table 4. Multinomial estimation of job choice, for 8-18 years old out of school

	Mean value variable	Agr. wage earner Relative risk	Agr. wage earner z of parameter	Non agr. wage earner Relative risk	Non agr. wage earner z of parameter
Individual characteristics					
Gender	0.74	6.88	11.3	0.38	-6.3
Indigenous	0.28	0.69	-2.1	0.69	-2.0
Completed grade					
No schooling	0.29	–	–	–	–
Incomplete primary	0.22	1.06	0.2	1.31	0.9
Primary 6	0.47	1.15	0.5	2.57	3.4
Secondary 1	0.01	2.19	1.0	6.24	2.3
Secondary 2	0.02	1.72	1.0	3.05	1.9
Secondary 3	0.19	1.60	1.5	6.25	5.5
Higher than secondary	0.01	2.50	1.1	4.74	1.9
Household characteristics					
Father not at home	0.10	2.05	3.1	0.99	0.0
Household head's education	2.09	0.89	-3.5	0.93	-2.4
Household head's wage (pesos per month)	5.57	1.00	-3.4	1.00	-3.7
Household head is ag. wage earner	0.54	13.46	14.2	4.15	7.6
Household head is non ag. wage earner	0.09	2.75	3.0	12.94	8.0
Household head is self employed	0.37	–	–	–	–
Irrigated land per adult (ha)	0.03	1.51	0.8	1.76	1.0
Rainfed land per adult (ha)	0.45	0.89	-2.3	0.81	-3.2
Number of children	3.64	1.04	0.8	1.15	3.0
Distance to urban center (in km)	101.0	1.01	5.8	1.01	3.3
State control					
Guerrero	0.08	–	–	–	–
Hidalgo	0.17	5.06	5.7	8.60	7.1
Michoacan	0.09	2.68	3.3	4.23	4.5
Puebla	0.19	13.81	8.5	17.94	8.8
Queretaro	0.06	1.16	0.4	11.44	7.1
San Luis Potosi	0.16	2.87	4.0	7.66	7.3
Veracruz	0.25	5.50	6.5	6.79	6.7
Number of observations	3620	2087		1241	
Pseudo R2	0.33				

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.

For many young people of poor villages, the better job option is, however, to leave their villages (in particular these poor villages that are the target of Progresa) and move to more active centers of population. 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 for work (and not to study or to get married). We estimated the probability of migrating for three age cohorts. The results are reported in Appendix Table 1, and a summary of the overall probability to migrate between 13 and 25 years old is reported in Table 5.

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

	Completed grade				
	Primary 6	Secondary 1	Secondary 2	Secondary 3	Post-secondary
By gender					
Boys	29.8	33.4	33.2	35.2	32.7
Girls	24.9	26.7	26.7	28.8	23.1
By occupation of head of household					
Ag. Worker	29.6	32.8	33.2	35.0	31.8
Non-age worker	24.5	26.1	26.5	28.8	20.0
Self-employed	25.3	28.6	27.8	30.5	26.0
By head of household education level					
Primary	27.9	31.1	31.0	33.3	30.0
More than primary	16.6	20.0	19.3	20.0	16.3

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

To construct life time earnings, we estimated earnings in each of these occupations as a function of gender, indigenity, education, and age, controlling for distance to an urban center and state. The estimation is done on the working population 8 years and older. The results are reported in Table 5. Gender difference is important with men earning 24% more than women in self-employed jobs, 39% more in agricultural wage and 40% in non-agricultural wage activities. Return to education are most important in non-agricultural wage and self-employed activities, although the range from the lowest education level to the highest remains less than a factor of 2 in earnings. This reveals that even activities other than agricultural wage work offer low income in these poor villages. This is partly due to the strong selectivity in the decision to remain in the villages. Similarly life time earnings for migrants need to be estimated. For that estimation, we resorted to the a different source of information, the Household Income and Expenditure Survey (ENIGH), which is based on a random sample of population, to estimate income earnings. This earning equation is reported in Table 7. The sample consists of those individuals 18 and older and the estimation is corrected for the endogenous choice of working. As with the other wage equations, the gender difference is again very significant, as males earn 34% more than females. The return to education is also very pronounced. The difference between the lowest return and the highest return is over 160%. Figure 2 summarizes the life cycle of earnings in the 4 activities.

Table 6. Wage equations
(Endogenous variable is log of monthy wage)

	Self-employed			Agricultural wage earner			Non agr. wage earner		
	Mean value of variable	Coef.	z	Mean value of variable	Coef.	z	Mean value of variable	Coef.	z
Gender	0.70	0.24	5.0	1.0	0.39	7.5	0.6	0.40	16.1
Indigenous	0.36	-0.28	-8.0	0.4	-0.16	-18.7	0.2	0.03	1.1
Age	42.3	0.016	0.9	36.8	0.014	3.8	28.5	0.025	1.6
Age^2 (/1000)		-0.28	-0.7		-0.25	-3.1		-0.32	-0.7
Age ^3 (/1000000)		0.30	0.1		1.02	1.8		0.57	0.2
Completed grade									
No schooling	0.25	–	–	0.23	–	–	0.07	–	–
Incomplete primary	0.46	0.01	0.2	0.43	0.01	0.8	0.27	0.10	2.1
Primary 6	0.18	0.08	1.5	0.23	0.02	1.5	0.35	0.12	2.5
Secondary 1	0.00	0.52	2.3	0.01	0.05	0.9	0.01	0.12	1.1
Secondary 2	0.01	0.36	2.5	0.01	0.03	0.5	0.02	0.22	2.7
Secondary 3	0.07	0.27	3.3	0.08	-0.04	-1.8	0.19	0.13	2.4
Higher than secondary	0.02	0.48	4.6	0.01	-0.05	-1.3	0.09	0.55	7.2
Distance to an urban center (in 1000 kms)	0.101	-1.47	-3.4	0.111	0.21	1.6	0.094	0.74	2.3
State control									
Hidalgo	0.14	1.13	14.2	0.18	0.07	2.8	0.18	0.08	1.3
Michoacan	0.08	1.01	10.2	0.07	0.41	13.8	0.10	0.15	2.2
Puebla	0.19	0.74	9.2	0.18	0.09	3.4	0.16	-0.07	-1.1
Queretaro	0.04	1.27	11.1	0.02	0.29	7.0	0.16	-0.01	-0.2
San Luis Potosi	0.17	0.86	10.7	0.12	0.10	3.9	0.18	0.02	0.3
Veracruz	0.28	0.94	12.1	0.39	0.04	1.6	0.18	0.02	0.3
Inverse Mills Ratio	1.72	-0.21	-2.9	0.64	0.17	5.9	1.47	-0.25	-5.7
Intercept	1	5.09	16.6	1	5.59	54.0	1	5.66	26.7
Average wage		290			548			506	
Number of observations		3,484			12,662			3,946	
R-squared		0.20			0.13			0.22	

Table 7. Wage equation for migrant
(Endogenous variable is log of monthy wage)

	Mean value variable	Coef.	z
Wage			
Individual characteristics			
Gender	0.45	0.288	5.3
Age	39.72	0.033	2.1
Age^2 (/1000)		-1.2E-04	-0.3
Age ^3 (/1000000)		-3.1E-06	-1.2
Completed grade			
No schooling	0.16	–	–
Incomplete Primary	0.26	0.334	7.6
Primary	0.22	0.757	16.3
Secondary 1	0.02	0.844	8.3
Secondary 2	0.02	1.109	12.8
Secondary 3	0.18	1.157	23.6
Higher than secondary 3	0.15	1.658	33.4
Inverse Mills Ratio		0.165	0.057
Intercept	1.00	4.811	21.4
Selection equation (endogenous variable = work)			
Individual characteristics			
Gender		1.500	60.9
Age		-0.012	-14.3
Married		-0.129	-5.3
Family size		-0.038	-6.4
Number of kids ≤ 12		0.029	2.6
Wealth indicators			
House has electricity		-0.129	-2.4
House has dirt floor		0.144	3.7
Completed grade			
No schooling		–	–
Primary		0.108	2.8
Secondary 1		0.176	4.2
Secondary 2		-0.044	-0.5
Secondary 3		0.152	1.8
Secondary 4		0.259	5.6
Higher than secondary 4		0.564	11.7
Intercept		0.182	2.3
Mean wage		886.6	
Number of observation		14882	
Censored observation		6523	
Uncensored observation		8359	
Wald chi2(10)		1879.6	
Correlation between errors of the selection and failure equations		0.05	(st. err. .05)

Estimated from ENIGH, 1996

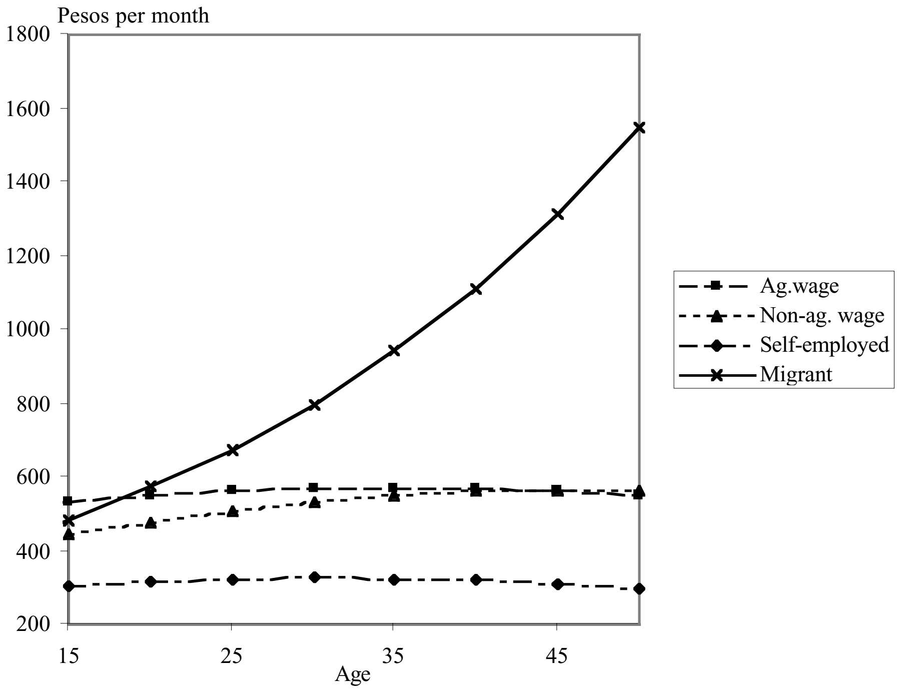

Figure 2. Life cycle of earnings in the 4 activities.

5.2. Failures at first try and upon repetition

Table 8 reports the estimation for the probability that a child fails a grade. The estimation is corrected for the endogenous choice of attending school. As the estimation of repeated failure turned out to be difficult because of small sample size and very similar to the estimation of first failure, we only report the joint estimation. 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 as 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, hence for the whole sample. Further work needs to be done to disentangle those different effects. It is interesting to note that once one controls for individual characteristics and for the decision to repeat a class first failed, the probability of failure is not different from the first time in the grade. Performance of the children is strongly affected by the presence of someone in the household with a higher education level, pointing either to the importance of a supportive environment. Interestingly the performance of poor children is no worse than that of richer children, and the distance to secondary school, which has a determining impact on enrollment, does not affect the performance. Note that the Progresa conditional transfers do not seem to influence the probability of failure.

Nothing really surprising in the selection equation, which is a reduced form of the model that will be estimated later. None of the three identification variables (house has a dirt floor, gender, and family size) are significant when introduced in the failure equation and jointly, we cannot reject their non-significativity (Chi2(3) = 4.94, p-value=.18).

Table 8. Probability of failing grades in secondary and post-secondary school, school year 1998-99

	Mean value variable	Coef.	z
Failure			
Individual characteristics			
Age–Grade+6	-0.353	0.476	3.8
(Age–Grade+6)*Grade	-2.591	-0.086	-4.7
Repeating	0.120	0.032	0.5
Transfer received (pesos/month)	0.5	-0.143	-1.3
Attended grade			
Secondary 1	0.400	0.208	3.6
Secondary 2	0.277	-0.102	-1.5
Secondary 3	0.240	–	–
Post-secondary	0.083	0.355	4.2
Household characteristics			
Max. education among other members	6.20	-0.019	-2.8
House has bathroom (0/1)	0.70	-0.051	-1.1
Poor	0.79	0.080	0.9
Progresa village	0.64	0.183	1.9
Distance to secondary school (km)	1.93	0.013	1.2
Intercept	1	-1.485	-13.3
Selection equation (endogenous variable = attending school)			
Individual characteristics			
Gender (1 = male)	0.531	0.130	3.5
Age–Grade+6	-0.232	-0.542	-5.0
(Age–Grade+6)*Grade	-1.867	0.041	2.5
Failed in 1997-98	0.143	-0.414	-8.1
Transfer received (pesos/month)	0.5	0.217	2.2
Attended grade			
Secondary 1	0.442	-1.104	-16.0
Secondary 2	0.228	0.088	1.0
Secondary 3	0.198	–	–
Post-secondary	0.132	-1.997	-24.3
Household characteristics			
Household head's education	3.00	0.078	9.4
Family size	7.1	-0.010	-1.2
House has bathroom	0.69	0.144	3.4
House has dirt floor	0.539	0.094	2.3
Poor	0.79	-0.153	-2.0
Progresa village	0.62	0.085	1.0
Distance to secondary school	2.02	-0.064	-6.4
Distance to urban center (in km)	107.28	0.003	7.1
Intercept	1	1.249	10.2
Number of observation		7825	
Censored observation		1597	
Uncensored observation		6228	
Percent failure (among attending school)		25.6	
Wald chi2(10)		126.13	
Correlation between errors of the selection and failure equations		0.76	(st. err. 0.08)

Standard errors adjusted for clustering on household
The reference group for the selection equation include all kids attending school in 1997.

5.3. Joint estimation of repeating and enrolling

Table 9 presents the estimation of the dynamic decision process on enrollment. Given the small number of children who 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 among these samples (and similarly 85% for Secondary 3). Despite these data limitations, some insights can be gained from the estimation results. Progresa has a positive influence for repeating the first year of secondary school. Since the projected life cycle wage measures the opportunity cost of repeating, it negatively affects enrollment at least in the first two years of secondary school. The future benefit of an additional year of schooling has a positive impact.

For the first-time enrollment decision, the estimation fares considerably better. The proxies for income, namely the dwelling characteristics, are significant in most of the regressions and show the appropriate signs. Family characteristics such as education of the household head, age of the mother, and presence of the father are also important determinants. Interestingly, however, the gender of the child is only significant in the decision to enroll in the first year of each cycle, where boys are more likely to enroll than girls The distance variables appear to be important determinants. As one would suspect, the distance to secondary school is an important school cost and it negatively affects the decision to enroll. Conversely, the distance from the capital has a positive influence since it most likely proxies out of school opportunities. Progresa has a positive impact for enrollment in the first and third year of secondary school. Progresa plays a peculiar role in the second year of secondar, by adversely affecting first-time enrollment. As in the decision to repeat, the wage variable again has a negative impact, while the future benefit of an additional year of school has a positive effect at each grade level.

5. Simulations and conclusion

Table 9. Enrollment decision in secondary and post-secondary grades, school year 1998-99

	Post-secondary school			Secondary 3			Secondary 2			Secondary 1		
	Mean value variable	Marginal effect	z (of coeff.)	Mean value variable	Marginal effect	z (of coeff.)	Mean value variable	Marginal effect	z (of coeff.)	Mean value variable	Marginal effect	z (of coeff.)
Enrollment decision first time												
Observed percentage enrollment	44.8			96.2			96.3			76.0		
Cost												
Gender	0.55	0.131	3.2	0.55	-0.004	-0.4	0.53	0.004	0.4	0.51	0.109	5.9
Indigenous	0.28	0.057	1.8	0.30	-0.005	-0.5	0.30	0.008	0.9	0.29	0.049	3.1
Household head's education	3.27	0.031	5.2	3.26	-0.002	-0.9	3.26	0.003	1.8	2.91	0.015	4.8
Rank among children	1.63	0.006	0.3	1.83	-0.003	-0.6	2.05	0.003	0.9			
Father not at home	0.10	0.034	0.8	0.09	0.002	0.1	0.10	-0.043	-3.8	0.09	0.057	2.7
Mother's age	40.51	-0.004	-2.3	40.03	0.000	-0.5	39.21	0.000	0.1	38.67	-0.002	-2.6
House has bathroom	0.73	0.097	3.2	0.73	0.018	2.1	0.70	-0.007	-0.6	0.64	0.032	2.2
Number of rooms in house	2.20	0.033	2.8	2.10	-0.008	-2.7	2.06	0.005	1.1	1.96	0.010	1.8
House has electricity	0.85	0.081	2.2	0.83	0.006	0.6	0.82	-0.004	-0.5	0.76	0.052	3.6
House has potable water	0.43	0.037	1.3	0.41	-0.007	-0.9	0.41	0.003	0.4	0.37	0.034	2.7
Poor	0.75	0.002	0.07	0.80	-0.026	-2.04	0.81	0.059	2.03	0.85	-0.043	-1.69
Distance to secondary school				1.80	0.001	0.5	1.88	-0.004	-2.3	2.25	-0.035	-10.8
Distance to capital	13.08	0.047	5.5	7.45	0.041	2.5	11.19	0.029	1.9	8.90	0.098	4.4
Tele-secondary school				0.85	0.020	1.8	0.85	-0.018	-1.6	0.83	0.049	3.0
Progresa village	0.59	0.033	1.3	0.62	-0.028	-1.8	0.65	0.061	2.3	0.61	0.002	0.1
Transfer				0.50	0.049	2.54	0.54	-0.060	-1.97	0.52	0.084	2.47
Net benefit												
Future benefits	13.08	0.047	5.5	7.45	0.041	2.5	11.19	0.029	1.9	8.90	0.098	4.4
Wage with current schooling	7.20	-0.348	-5.0	7.15	-0.029	-1.5	7.03	-0.054	-2.0	6.77	-0.185	-5.2
Number of observations		1724			2478			3115			5206	
Pseudo R2		0.07			0.04			0.06			0.11	
Repetition, in case of failure												
Observed percentage enrollment	42.6			84.1			89.4			49.4		
Cost												
Gender	0.54	0.072	0.5	0.50	-0.083	-1.4	0.51	0.023	0.7	0.50	0.204	3.3
Indigenous	0.24	0.112	1.4	0.25	0.142	3.0	0.29	0.012	0.4	0.28	-0.093	-1.1
Rank in children	1.48	0.051	0.9	1.74	0.000	0.0	1.97	0.020	1.5	2.06	0.042	1.8
Family size	6.79	-0.010	-0.6	6.92	0.003	0.3	7.16	0.001	0.2	7.43	0.010	1.0
Father not at home	0.12	-0.064	-0.6	0.09	-0.005	-0.1	0.12	0.003	0.1	0.11	0.265	2.5
Household head's education	3.54	0.045	3.0	3.33	0.016	1.7	2.90	0.001	0.3	2.70	-0.010	-0.6
House has bathroom	0.74	0.140	1.9	0.73	0.024	0.5	0.69	-0.059	-1.8	0.60	-0.064	-1.0
Number of rooms in house	2.38	0.002	0.1	2.20	0.007	0.5	2.00	0.015	1.1	1.96	0.002	0.1
House has electricity	0.87	0.059	0.6	0.84	0.024	0.5	0.82	-0.012	-0.4	0.78	0.010	0.2
House has potable water	0.49	-0.035	-0.5	0.48	0.022	0.6	0.43	-0.020	-0.9	0.43	0.069	1.6
Poor	0.66	-0.081	-1.1	0.75	-0.015	-0.2	0.77	0.308	2.2	0.82	-0.141	-1.6
Distance to secondary school				1.68	-0.004	-0.3	1.95	-0.011	-2.0	2.34	0.006	0.4
Tele-secondary school				0.87	0.030	0.5	0.88	-0.001	0.0	0.81	-0.029	-0.5
Distance to capital	153.10	0.001	2.0									
Progresa village	0.63	0.068	1.0	0.61	-0.095	-1.4	0.62	0.125	1.3	0.62	-0.134	-1.3
Transfer				0.5	0.082	1.0	0.5	-0.194	-1.5	0.5	0.251	2.4
Net benefit												
Future benefits/success	4.83	1.217	0.8	6.56	0.107	0.6	10.85	0.077	1.7	7.63	0.439	2.3
Future benefits/failure	2.61	1.407	0.8	0.62	0.494	2.7	0.45	0.448	3.9	1.22	0.181	0.8
Wage with current schooling	7.18	-1.400	-0.8	7.11	-0.062	-0.4	7.05	-0.224	-2.8	6.80	-0.691	-3.6
Number of observations		284			320			303			778	
R2		0.08			0.17			0.28			0.14	

References

Acemoglu, Daron, J.-S. Pischke. “Changes in the wage structure, family income, and children’s education.” *European Economic Review* 45(2001) 890-904

Altonji, Joseph G.. “The Demand for and Return to Education When Education Outcomes are Uncertain.” *Journal of Labor Economics*, Volume 11, Issue 1, Part1:Essays in Honor of Jacob Mincer (Jan. 1993), 48-83.

Appleton, Simon, John Hoddinott and John Knight. “Primary Education As An Input Into Post-Primary Education: A Neglected Benefit.” *Oxford Bulletin of Economics and Statistics*, 58,1 (1996) 211-219.

Behrman and Todd, 1999.

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

Cameron, Stephen V., James J. Heckman. “The Dynamics of Educational Attainment for Blacks, Hispanics, and White Males.” *The Journal of Political Economy*, 2001, Volume 109, Number 3, 455-499.

Cameron, Steven, James J. Heckman “Life Cycle Schooling and Dynamic Selection Bias: Models and Evidence for Five Cohort of American Males.” *The Journal of Political Economy*, 1998, Volume 106, Issue 2, 263-333.

Card, David E. and Thomas Lemieux. “Adapting to circumstances : the evolution of work, school, and living arrangements among North American youth”. National Bureau of Economic Research, 1997, WP no. 6142.

Card, David E.; Krueger, AB. “Does School Quality Matter - Returns to Education and the Characteristics of Public Schools in the United-States.” *Journal of Political Economy*, 1992 Feb, V100 N1:1-40

Card, David E.; Krueger, AB. “School Resources and Student Outcomes: An Overview of the Literature and New Evidence from North and South Carolina.” *Journal of Economic Perspectives*, 1996 Fall, V10 N4:31-50.

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

Comay, Y., A Melnik, M.A. Pollatschek. “ “.*International Economic Review*, Volume 14, Issue 2(June 1973), 421-435.

Eckstein, Zvi, Kenneth I. Wolpin.”Why Youths Drop Out of High School: The Impact of Preferences, Opportunities, and Abilities.” *Econometrica*, Vol 67, No.6(Nov 1999) 1295-1339

Hanushek, Eric. “Interpreting Recent Research on Schooling in Developing Countries.” *World Bank Research Observer* (1996), 227-246

Haveman, Robert and Barbara Wolfe. “The Determinants of Children’s Attainments: A Review of Methods and Findings.” *Journal of Economic Literature*, Volume 33, Issue 4 (Dec., 1995), 1829-1878

Kane, Thomas J. “College Entry by Blacks since 1970: The Role of College Costs, Family Background, and the Returns to Education.” *Journal of Political Economy*, Volume 102, Issue 5(Oct. 1994), 878-911

Keane, Michael P., Kenneth I. Wolpin. “The Career Decisions of Young Men.” *The Journal of Political Economy*, Volume 105, Issue 3(June 1997), 473-522

Neumark and Wascher (1998). “Minimum-Wage Effects on School and Work Transitions of Teenagers.” *American Economic Review*, 1995 May, V85 N2:244-249

Rees, Daniel I. and H. Naci Mocan. “Labor Market Conditions and the High School Dropout Rate: Evidence from New York State.” *Economics of Education Review*, Volume 16, No. 2, (1997) 103-109

Ribar, David C.. "The effects of local employment opportunities on youth's work and schooling." *Economics of Education Review*, Volume 20 (2001), 401-413

Rochat, Denis, Jean-Luc Demeulemeester. "Rational choice under unequal constraints: the example of Belgian higher education." *Economics of Education Review* 20 (2001) 15-26

Skoufias, Emmanuel, Benjamin Davis, and Jehre Behrman, 1999

Willis, Robert J., Sherwin Rosen. "Education and Self-Selection." *The Journal of Political Economy*, Volume 87, Issue 5, Part 2: Education and Income Distribution (Oct. 1979), S7-S36.

Yunez-Naude, Antonio and J. Edward Taylor. "The Determinants of Nonfarm Activities and Incomes of Rural Households in Mexico, with Emphasis on Education." *World Development*, Vol. 29, No. 3, pp. 561-572.

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									
Gender	0.52	0.29	1.0	0.51	1.68	3.9	0.49	1.18	4.1
Age in 1997	14.5	1.17	7.8	18.4	0.31	1.6	23.0	-0.15	-1.5
Indigenous	0.29	-0.39	-1.1	0.28	0.44	0.8	0.29	-0.31	-0.9
Completed grade									
No education	0.04	–	–	0.08	–	–	0.11	–	–
Incomplete primary	0.24	0.93	0.9	0.21	6.77	4.0	0.27	0.47	0.8
Primary 6	0.36	1.71	1.8	0.38	7.12	4.9	0.39	1.31	2.1
Secondary 1	0.12	1.21	1.1	0.02	12.17	3.7	0.01	3.37	1.8
Secondary 2	0.12	2.39	2.1	0.03	9.07	3.5	0.02	3.00	2.0
Secondary 3	0.10	2.50	2.1	0.20	11.72	6.1	0.16	2.40	2.9
Higher than secondary	0.02	-0.38	-0.3	0.08	13.55	5.7	0.04	3.16	2.7
Household characteristics									
Father not at home	0.16	1.33	2.9	0.00	0.00	0.0	0.00	0.00	0.0
Household head's education	2.47	-0.18	-2.6	2.41	-0.30	-3.1	3.07	-0.31	-5.0
Household head's wage (pesos per montl	575	0.00	-2.4	566	0.00	-2.0	601	0.00	-0.6
Household head is ag. wage earner	0.48	1.03	2.9	0.46	0.51	1.1	0.49	0.03	0.1
Household head is non ag. wage earner	0.09	0.41	0.6	0.08	1.02	1.1	0.09	0.28	0.5
Household head is self employed	0.44	–	–	0.46	–	–	0.42	–	–
Number of children	3.68	0.44	4.5	2.88	0.69	5.9	1.46	0.61	8.3
Irrigated land per adult (ha)	0.04	0.48	1.9	0.02	-0.42	-0.3	0.02	-0.17	-0.2
Rainfed land per adult (ha)	0.55	0.05	0.4	0.43	-0.30	-1.2	0.37	-0.22	-1.1
Distance to urban center (in km)	103.6	0.03	7.0	100.7	0.03	4.9	100.3	0.02	4.7
State control									
Guerrero	0.08	–	–	0.08	–	–	0.08	–	–
Hidalgo	0.17	6.43	4.7	0.17	6.52	4.0	0.17	3.25	2.8
Michoacan	0.13	8.44	5.2	0.12	10.02	5.0	0.12	7.47	4.6
Puebla	0.16	6.90	4.7	0.16	6.15	3.6	0.16	2.76	2.3
Queretaro	0.06	8.81	4.6	0.07	10.58	4.8	0.06	8.43	4.6
San Luis Potosi	0.16	6.06	4.3	0.16	8.12	4.7	0.16	4.07	3.3
Veracruz	0.25	8.62	6.3	0.24	8.43	5.3	0.24	4.74	4.1
Number of observations	13205			10122			9321		
Observed percentage migrating	4.12			5.87			3.01		
Pseudo R2		0.06			0.05			0.09	