IMF Working Paper

How Good Are Ex Ante
Program Evaluation Techniques?
The Case of School Enrollment in
PROGRESA

Fabian Bornhorst

INTERNATIONAL MONETARY FUND

IMF Working Paper

Fiscal Affairs Department

**How Good Are Ex Ante Program Evaluation Techniques?
The Case of School Enrollment in PROGRESA**

Prepared by Fabian Bornhorst[1]

Authorized for distribution by Mark Horton

September 2009

Abstract

This paper evaluates a microsimulation technique by comparing the simulated outcome of a program with its actual effect. The ex ante evaluation is carried out for a conditional cash transfer program, where poor households were given money if the children attended school. A model of occupational choice is used to simulate the expected impact of the program. The results suggest that the transfer would indeed increase school attendance and do more so among girls than boys. While the simulated effect tends to be larger than the actual effect, the latter lies within bootstrapped confidence intervals of the simulation.

Contents   **Page**

Tables

Figures

I. Introduction

Evaluation of social programs is particularly important in the context of poverty reduction in developing countries, where programs aiming to improve the human capital of the poor are being implemented at a large scale. This importance derives from the key role that is attributed to poverty reduction in the process of economic and social development and for reasons of political accountability.[2] Policymakers need tools to assess the impact of such programs, and ideally, evaluations should not only be carried out after completion of a project, but also prior to its implementation. Ex ante evaluation techniques using microsimulation can help to improve program design and avoid failure by providing quantifiable information about the direction and magnitude of effects that are likely to occur.[3]

While the use of microsimulation techniques is gaining popularity, little is known about their accuracy and an assessment is currently lacking. The evaluation of evaluation methods can be methodologically challenging, because, absent a counterfactual, the impact of a program itself is difficult to single out, let alone compare with predictions of a model.

This paper tries to fill this gap, and for a specific type of program, offers an evaluation of a microsimulation technique. The program used in this paper was implemented as a randomized experiment, and therefore lends itself as a suitable benchmark for the evaluation of the microsimulation method. The approach taken in this paper is similar to Lalonde (1986) who uses experimental data as a benchmark to evaluate the performance of the difference–in–differences estimator for differently constructed counterfactuals.

The incidence of child labor varies greatly across the world, with an estimated 16 million children between 5 and 14 (or 17 percent of the cohort) working in Latin America alone (ILO, 2003).[4] The nature of child labor differs between urban and rural areas (OECD, 2003).

---

[2] See, for example, the World Development Report 2000 on "Attacking Poverty" (Worldbank, 2000).

[3] Microsimulation is a general term for methods designed to simulate systems at the level of individual units rather than the overall population.

[4] A common distinction is made between child work and child labor (see, e.g. (Anker, 2000)). While child work comprises light activities that are not considered harmful to the educational opportunities of the child, all

(continued…)

In rural areas, children in poor households often contribute to household production in agriculture and livestock, or are engaged in looking after younger siblings, and as a result do not attend school. Lack of education, however, substantially narrows the future employment opportunities of children, and reduces their chances of breaking the poverty cycle.

School attendance, especially at secondary level, comes at a cost for households. Children who go to school cannot contribute to household income in the form of home production or labor income. In addition, even if education itself was free, school attendance is often associated with additional costs for transportation and schooling supplies, such as books. In the trade off between future and current income, poor households are often forced to decide against secondary school attendance for their children.

PROGRESA is a well known program in rural Mexico, which has been in place since 1997.[5] It aims at improving the educational attainment and health status of the rural poor. One of the main goals of the program is to increase secondary school enrollment ratios. To this end, eligible families are offered a cash transfer conditional on the school attendance of their children.

In this paper, data from a survey conducted before the introduction of the program are used to estimate a model of occupational choice and simulate the potential impact of PROGRESA on school enrollment ratios. The microsimulation exercise follows closely an existing model of occupational choice (Bourguignon et al., 2002). In this model, children decide between three occupational choices: they either work, work and attend school, or go to school only. The contribution of children to the income of a household through home production is explicitly accounted for. The simulations are then compared to the observed impact of the program.

---

activities that damage children's physical and psychological health and are detrimental to their future development, in particular education, are considered child labor.

[5] It is now called OPORTUNIDADES, the main features of PROGRESA are outlined in Appendix A. In 2001, total spending under this program roughly amounted to 0.2 percent of Mexican GDP (Attanasio et al., 2001).

The analysis shows that the predictions of the model come close to the observed effect. Moreover, disaggregated by age and gender, the observed effects are generally within bootstrapped confidence intervals of the simulation exercise. While the overall performance of the microsimulation is good, the simulation shows two shortcomings. First the model has a poor fit in subsamples where the three options considered in the multinomial choice model (school attendance only, work and school, and work only) are unequally distributed across the sample. Second, the model does capture well the transition to secondary school, as it predicts significantly more children to complete the transition than actually do.

The remainder of the paper is organized as follows. The next section discusses the model selection and how the model can be used for ex ante simulations. Section III presents the results for the simulation exercise. In Section IV results from a conventional ex post evaluation are presented and related to similar findings in the literature. In Section V, ex ante and ex post results are compared and some lessons from the simulation exercise are drawn. The last section concludes.

II. EX ANTE EVALUATION: THEORY

Choosing a model for a microsimulation exercise involves striking a balance between structural sophistication and feasibility. Structural models offer insights into individual behavior by identifying deep parameters that go beyond reduced form specifications. However, some parameters in structural models can be difficult to identify due to the lack of valid instruments, which reduces the empirical value of such models. Many aspects of a structural models, while interesting, may be irrelevant for the simulation exercise, as long as the reduced form model is general enough and implicitly accommodates these aspects. Reduced form models, on the other hand, need to exhibit a certain degree of structural sophistication that allows identification of those parameters that are relevant for the microsimulation.

Economic theory offers a range of models that would lend itself for the purpose of microsimulating the effect of PROGRESA, for example, models of household time allocation decisions, models of lifetime income considerations, and models that outline the effect of education on human capital. Attanasio et al. (2001) develop a structural model in which each

child maximizes his or hers lifetime earnings and takes schooling decisions accordingly, given a market wage and an intertemporal discount factor. In this context, the authors study the impact of varying levels for an educational grant, given some value for the discount factor. While appealing from a modeling point of view, one could argue that in rural Mexico the trade–off around which the model is built is often dominated by liquidity constraints: poor households may not have access to the financial intermediaries that could facilitate this intertemporal decision. Together with the possible lack of lifetime planning horizons in poor families, this could be seen as a significant limitation to the applicability of the model.

Todd and Wolpin (2003) develop a dynamic structural model at the household level. This approach addresses a major shortcoming of other models which optimize at the individual level by explicitly considering interactions between family members. In their model, parents maximize parental lifetime income by choosing an optimal fertility rate and deciding upon the time allocation of their children. Fewer but better educated children are seen as an alternative that becomes more attractive the lower the costs of education. This addresses more fundamental questions than the mere schooling decision and could be useful to assess the long term impact of a social program. In the case of evaluating PROGRESA, where the time horizon is limited to two years by data considerations, it is unlikely that variations in the fertility behavior can be observed or identified.

Finally, Bourguignon, Ferreira and Leite (2002) develop a model of a discrete labor supply decisions, where a child, contributing to the household income, decides either to go to school, to work, or to mix between the two options. The trichotomous framework is especially suitable for developing countries where many children do both, attend school and are engaged in some kind of economic activity, especially in rural areas (see for example OECD (2003, Table A11)). The model relates household income and children's earnings to the labor supply outcome for each child. The reduced form of the model also encompasses contributions of children to home production, an important but often neglected aspect that is prevalent among the rural poor. It is however silent with respect to interactions at the household level, or lifetime earnings considerations. The model was developed around a conditional cash transfer scheme, the *Bolsa Escola* program in Brazil, which shares basic characteristics with PROGRESA.

A. A Model of Occupational Choice

The model is based on the following assumptions.[6] The unit of decision is the child and not the household. Hence, considering the intrahousehold labor allocation, it is assumed that schooling decisions of children are taken independently from each other and labor market outcomes of other family members are unaffected. Furthermore, the composition of households is taken as exogenous, and effects on fertility are disregarded.

Let S_i be a choice variable that takes one of the values $j = 0, 1, 2$ for each child i

$$S_i = \begin{cases} 0 \text{ if } i \text{ does not attend school} \\ 1 \text{ if } i \text{ works and attends school} \\ 2 \text{ if } i \text{ attends school and does not work} \end{cases}$$

Under choice $j = 0$ it is assumed that individual i works full time, either being engaged in activities in the household or by earning a wage on the labor market. Under choice $j = 1$ it is assumed that the individual does both, attend school and participates in the formal labor market. Finally, under choice $j = 2$ it is assumed that children attend school and do not work for remuneration, however, they may allocate some time to household work in a way that will be described further below. In the framework of a multinomial choice model, each i will make an optimal choice according to:

$$S_i = k \text{ if } S_k(.) > S_j(.) \text{ for all } j \neq k$$

where $S_j(.)$ is some function $S_j(X_i, H_i, Y_{ij}, v_{ij})$ of the following variables:

X_i : characteristics of i (age, schooling etc.)
H_i : characteristics of the household of i (parental education etc.)
Y_{ij} : income of the household of i when choosing j
v_{ij} : i.i.d. shock

The household's income Y_{ij} is the sum of i's own income y_{ij} and the income of the rest of the household, Y_{-i}. After combining all non income variables into $Z_i = [X_i, H_i]$ and linearizing, the model takes the following random utility representation:

---

[6] This section draws mainly on Bourguignon et al. (2002).

$$(1) \qquad U_{ij} = \gamma_j Z_i + \alpha_j (Y_{-i} + y_{ij}) + v_{ij}$$

Even though this specification does not explicitly model the deep parameters determining schooling decisions, it contains the relevant variables through which the program operates, namely the household income (which is directly affected by the conditional transfer) and the income a child could achieve by participating in the labor market. Note also that the coefficients γ_j and α_j may vary with each choice.

According to the individual's characteristics and prevailing labor market conditions, each individual can be attributed a measure of *potential* market earnings. Assume for the moment that the potential earnings of i are observable and denote this amount by w_i. Then, depending on the amount of time dedicated to work, the actual contribution of i's income y_{ij} under the three options is:

$$
\begin{aligned}
y_{i0} &= Kw_i \\
(2) \qquad y_{i1} &= My_{i0} = MKw_i \\
y_{i2} &= Dy_{i0} = DKw_i
\end{aligned}
$$

These expressions show that under choice $j = 0$ (not to attend school) a fraction K of the potential earnings is realized. Under choice $j = 1$ (attend school and work), only a fraction MK of w_i is realized. If i goes to school ($j = 2$) it may still contribute to home production for a value of DK times the potential market earnings. Combining (2) with (1) yields:

$$
\begin{aligned}
U_{i0} &= \gamma_0 Z_i + \alpha_0 Y_{-i} + \beta_0 w_{i0} + v_{i0} \\
(3) \qquad U_{i1} &= \gamma_1 Z_i + \alpha_1 Y_{-i} + \beta_1 w_{i1} + v_{i1} \\
U_{i2} &= \gamma_2 Z_i + \alpha_2 Y_{-i} + \beta_2 w_{i2} + v_{i2} \\
&\text{with } \beta_0 = \alpha_0 K, \ \beta_1 = \alpha_1 MK, \ \beta_2 = \alpha_2 DK.
\end{aligned}
$$

If potential earnings could be observed for each i and estimates of the coefficients α_i were available, the model in (3) could be used for microsimulation by looking at the effect of an exogenous variation of the household income under the schooling options $j = 1, \ 2$.

B. Estimation and Identification

Earnings can only be observed for those children who work for remuneration. While home production may contribute substantially to the household's non-cash income, it remains

unobservable. However, as discussed above, the representation in (3) requires some value for w_i for each individual no matter if the income is realized in the labor market or through home production. A solution is to use the observed market earnings under choices $S_i = \{0, 1\}$ and impute *potential* earnings for those individuals that either work at home or go to school. Following the literature, the observed earnings w_i can be modeled by:

$$(4) \qquad \log w_i \;=\; \delta X_i + m \cdot \mathbf{1}\{S_i = 1\} + u_i.$$

The vector X_i contains the typical regressors for a wage equation, in particular age, years of schooling, etc. The indicator function $\mathbf{1}\{...\}$ accounts for the fact that earnings of a child under option $S_i = 1$ could be lower because of school attendance. The coefficient m provides an estimate of the structural parameter M, which indicates the difference in earnings for those individuals who go to school and work. A control for sample selection (e.g. Heckman (1979)) would be necessary to control for possible selection biases. For those individuals where no earnings are observed, a potential wage $\tilde{w}_i$ can be imputed using the coefficient estimates from (4) and a random draw from the residuals vector $\hat{u}_i$.

Assuming exponentially distributed errors, the choice model (3) is known as the multinomial logit model (McFadden, 1973). In this model, the coefficients are identified only relative to a certain choice category. In the following, choice $j = 0$ is chosen as the base category. In this case, the multinomial choice model would yield estimates of the relative coefficients $g_k = (\gamma_k - \gamma_0)$, $a_k = (\alpha_k - \alpha_0)$ and $b_k = (\beta_k - \beta_0)$ for $k = 1,\ 2$. While in most applications this is sufficient, in the present case this level of identification is insufficient. Since the cash transfer is state dependent, it is necessary to identify all three coefficients $(\alpha_0,\ \alpha_1,\ \alpha_2)$ related to the household income.

As argued in Bourguignon et al. (2002) , the coefficients can be identified by making a simple structural assumption on time allocation K. Call the estimated coefficients from the multinomial logit model $\hat{a}_1$ and $\hat{a}_2$. Then it follows that

$$\alpha_1 - \alpha_2 = \hat{a}_1$$
$$\alpha_2 - \alpha_0 = \hat{a}_2$$
$$\alpha_1 MK - \alpha_0 K = \hat{b}_1$$
$$\alpha_2 DK - \alpha_0 K = \hat{b}_2$$

The estimate of the coefficient of the indicator function in equation (4) yields an estimate of M, $\hat{M} = \exp^{\hat{m}}$, because this is an estimate of the difference in earnings realized under alternative $j = 1$. With an estimate of M, any arbitrary combination of K and D allows to identify the coefficients α:

$$\hat{\alpha}_1 = \frac{\hat{a}_1 - \hat{b}_1}{1 - \hat{M}}$$
$$\hat{\alpha}_0 = \hat{\alpha}_1 - \hat{a}_1$$
$$\hat{\alpha}_2 = \hat{\alpha}_1 + \hat{a}_2 - \hat{a}_1$$
$$\hat{D} = \frac{\hat{b}_2 / K + \hat{\alpha}_0}{\hat{\alpha}_2}$$

Assume that $K = 1$, i.e. children who do not attend school realize their full potential earnings either in the labor market or through home production. For those individuals for which earnings are observed, this means that this is the only source of contribution to household income.

C. Impact Simulation

Having identified the levels of the income coefficients α_j the impact of the cash transfer is simulated using the following conditional payment:

$$U_{i0} = \gamma_0 Z_i + \alpha_0 Y_{-i} + \beta_0 w_{i0} + v_{i0}$$
$$U_{i1} = \gamma_1 Z_i + \alpha_1 (Y_{-i} + TR_i) + \beta_1 w_{i1} + v_{i1}$$
$$U_{i2} = \gamma_2 Z_i + \alpha_2 (Y_{-i} + TR_i) + \beta_2 w_{i2} + v_{i2}$$

where TR_i is the transfer paid conditional on school enrollment, which in turn depends on i's characteristics.

To complete the simulation it is necessary to obtain estimates for the residuals v_{ij}. In multinomial choice models, the residual terms $v_{ik} = (v_{ik} - v_{i0})$ can neither be observed nor precisely estimated. However, for each i, the set of residuals v_{i0}, v_{i1} and v_{i2} are bound to

belong to certain intervals, such that given the parameter estimates and i's characteristic they are *consistent* with the actual choice. In particular, if the choice $j = 0$ is defined as a base choice (meaning its utility is normalized to 0) and individual i has chosen choice $j = 1$, it must hold that:

$$U_{i1} - U_{i0} > 0$$
$$\text{and } U_{i1} - U_{i2} > 0$$
$$Z_i(\gamma_1 - \gamma_0) + Y_{-i}(\alpha_1 - \alpha_0) + y_{ij}(\beta_1 - \beta_0) + (v_{i1} - v_{i0}) >$$
$$\text{SUP}[0, Z_i(\gamma_2 - \gamma_0) + Y_{-i}(\alpha_2 - \alpha_0) + y_{ij}(\beta_2 - \beta_0) + (v_{i2} - v_{i0})]$$

where SUP denotes the supremum of the expression, and 0 is the normalized utility level of choice $j = 0$, and a similar expression holds for choices $j = 1, 2$.

These expressions place conditions on each pair of errors $\{(v_{i1} - v_{i0}), (v_{i2} - v_{i0})\}$, each of which is drawn from a double exponential distribution, yielding errors that are consistent with the choice and the assumed error distribution of the multinomial logit model. For each i a set of consistent random errors can be drawn from the inverse of the cumulative distribution of the residuals conditional on observing a choice k.[7]

III. EX ANTE EVALUATION: RESULTS

Data for PROGRESA was collected in five waves between November 1997 and November 1999.[8] The first two waves of data are the pre–program survey data and were collected in November 1997 and in March 1998. However, only the first wave asked income related questions and questions regarding the occupational choice. Hence, only this wave will be used to perform the ex ante simulation. The focus group are children between 10 and 16 years old. As Schultz (2001) points out, the school enrollment of the 6 to 9 year olds is almost 100 percent.

In the first wave, 17.5 percent of the respondents between 10 and 16 years old reported to receive payment for work. Table 1 summarizes the weekly wage by age. To put these

---

[7] For further details, see Appendix C.

[8] For a detailed description of how the variables are derived from the survey questionnaire see Appendix B.

numbers in perspective, the household income per capita is also reported. It shows that children who work make a significant contribution to the household's cash income, and this contribution increases with age.[9]

Table 1: Weekly Earnings and Per Capita Household Income

| Age | Obs | Average | | Median | |
		Earnings	Per capita income	Earnings	Per capita income
10	572	33	42	27	30
11	550	38	42	27	30
12	423	52	43	27	31
13	379	70	46	39	32
14	450	111	47	100	34
15	704	149	53	120	36
16	878	152	58	140	41
10-16	3,956	95	47	60	33

Note: Table reports earnings in Mexican pesos for those children that report to work for remuneration, and income per capita based on total household earnings. 10 pesos roughly corresponded to USD 1.10 in that period.

A breakdown by gender shows that girls are less likely to attend school throughout all age groups (Table 2). On average, 26.2 percent of the boys and 31.2 percent of the girls do not attend school. Enrollment ratios drop significantly after the age of 11. It is also evident that the option "work and go to school" is particularly predominant in younger age groups.

Table 3 reports sample means of some variables by occupational choice. On average, children who do not go to school or work and go to school come from households with higher income, supporting the previous finding that their work contributes significantly to the endowment of the household. Those who attend school are, on average, younger and have a higher rank.[10] Those who do not go to school tend to come from states with a higher median

---

[9] For Brazil, Bourguignon et al. (2002) find similar values for individual earnings, however, the relative contribution to household income is somewhat lower in their case. It could be that households in rural Mexico tend to be rather large and include many (inactive) members, reducing the the per capita figures.

[10] The rank is defined as the position of the child with respect to all household members below the age of 19. For example, a child with rank 3 has 2 elder siblings, and a child with rank 1 is the oldest child in the household below 19.

wage, although there is considerable variation across gender and age not reported in the table. Children who go to school tend to come from households with higher parental education.

Table 2: Reported Status

	Age							
	10	11	12	13	14	15	16	10-16
Boys								
not school	2.5	4.4	11.1	21.7	34.3	51.3	66.8	26.2
work and school	17.2	18.4	13.6	10.2	7.6	6.3	5.6	11.5
school only	80.3	77.2	75.3	68.1	58.1	42.5	27.6	62.3
Observations	1,818	1,667	1,806	1,665	1,628	1,623	1,477	11,684
Girls								
not school	3.0	4.4	17.4	30.5	44.6	60.4	73.2	31.2
work and school	16.8	17.4	10.2	6.7	3.4	2.4	2.0	8.9
school only	80.2	78.2	72.4	62.8	52.0	37.2	24.8	59.9
Observations	1,724	1,700	1,611	1,581	1,490	1,493	1,293	10,892

Note: Values indicated the percentage share for each cell. For a definition of status, see text.

Table 3: Sample Means

Variable	Not school	Work and school	School only
age	14.45	11.85	12.28
female	0.52	0.41	0.47
years of schooling	5.74	4.61	5.06
primary school completed	0.28	0.83	0.61
rank	1.73	2.13	2.19
children below 6 in household	0.87	0.94	0.88
median child wage in state	90.0	36.2	41.7
number of total hh members	7.31	6.96	7.2
weekly per capita household income	53.93	53.21	42.56
years of schooling of most educated parent	2.65	3.18	3.39
age of the oldest parent	46.27	43.93	45.03
share of control	0.40	0.34	0.38
Observations	6,478	2,306	13,795

Note: Data from children 10-16 years old in the wave 1 that were included in the estimation.

There is significant variation of sample means across age and gender, which is especially relevant in the case of PROGRESA, where the transfer scheme is grade-dependent and differentiates between boys and girls. Even though the data set is very rich with 22,576 observations, a breakdown by gender, age, and occupational status would lead to some cells containing relatively few observations. The estimations reported in the following section take

this into account by reporting results for age groups. The 10-11 years old form the first group, the 12-14 years old the second and the 15-16 years old constitute the last group.

A. Estimation of the Earnings Vector

The labor market earnings of each child, potential or realized, are an important aspect of the decision to attend school or not. However, as mentioned above, market wages are only observed for those children that report to work for remuneration and have to be imputed for all other individuals.

This section reports the estimates of actual and potential earnings as described in equation (4). The treatment of sample selection biases can be complex in the case of a multinomial choice model.[11] In the present case modeling the binary decision to work or not work (while controlling for the fact that some attend school *and* work), and using a two step Heckman estimation to correct for sample selection, yields satisfactory results.[12]

Results for one age group are presented in Table 4. Note that two exclusion restrictions are included in the selection equation. The rank of the child and the presence of the father in the household are assumed to influence the decision to work or not to work only, but not to influence the wage. The higher the rank of a child the lower is the probability of working. This can be seen as evidence for the widely held view that first-born children are sometimes disadvantaged with respect to school attendance. The absence of a father in the household increases the likelihood that children have to work in the labor market. Having completed primary education also increases the probability of participating in the labor market. Finally, girls are less likely to be wage earners than boys.

---

[11] For a discussion of selection biases in the context of multinomial choice models (Lee (1983)), see Bourguignon et al. (2001a).

[12] In contrast to Bourguignon et al. (2002) in the case of PROGRESA sample selection is found to be significant. Not accounting for this bias would lead to inconsistent estimates of the coefficients, which, in turn, would give biased estimates of the potential earnings for those individuals that do not report earnings.

In the estimation of the wage equation, the variables have the expected effect. Females earn significantly less than males, and, as expected, those who work and go to school also earn significantly less compared to those that just work. One important variable that accounts for regional variations is the median earnings of children of that particular age/gender group in the respective state. This variable, which proves to be significant, captures age and regional specific circumstances of the labor market. Additional years of schooling increase the wage, however, the effect is not significant in every age group since the degree of schooling varies little within the group.

Table 4: Estimation of Earnings Equation

Log(earnings)	Coefficient	t-value
female	-0.30	-4.27
years of schooling	0.059	0.68
(years of schooling)^2	-0.009	-0.92
median state earnings	0.25	5.36
status (1 if $S_j=1$)	-0.99	-23.21
Selection equation		
rank	-0.08	-4.08
father in hh	-0.26	-5.70
female	-0.37	-10.65
primary school completed	0.17	2.89
years of schooling	0.45	8.73
(years of schooling)^2	-0.05	-9.66
median state earnings	0.17	5.64
$\hat{\rho}$	0.75	
$\hat{\sigma}$	0.84	
$\hat{\lambda}$	0.63	3.49
Number of obs.	9,109	
Censored obs.	7,920	
Uncensored obs.	1,189	

Table presents estimates for equation 4 for using a Heckman two step estimation to correct for sample selection. Age group 12-14 years old. Constant included but not reported. Bottom and top one percentile of the wage distribution were not included.

To impute potential wages for those individuals who do not work, earnings were estimated using the age specific coefficients and adding a random draw from the estimated residuals vector obtained from the earnings equation. With a complete earnings vector $\hat{w}_i$ at hand for all i, one can proceed to estimate the choice model.

As a result of the sample selection bias, the observed and imputed earnings would not necessarily be expected to be the same. For example, if an individual has a lower imputed wage compared to the average wage observed among his peer group, this will have contributed to his schooling/working decision. Table 5 compares, by age group, the imputed median earnings with the observed earnings. While in the youngest age group the imputed earnings are higher than the observed ones, the reverse is true for the other age groups.

Table 5: Actual and Imputed Earnings

	Age group			
	10-11	12-14	15-16	10-16
Imputed wage	47	30	95	45
Observed wage	27	45	125	60

Note: Table shows estimates of the median imputed and actual weekly earnings. Bottom and top 1 percentile for each age group was dropped.

B. Estimation of the Choice Model

The choice model was also fitted separately by age group, catering for the fact that choices are unequally distributed among age groups. For example, for the very young the proportion of choice $S_i = 0$ (not school) is very small, whereas for the older children the option $S_i = 1$ (school and work) only has a small share (Table 2). Coefficients and t-values of the estimation of the multinomial choice model for one age group are reported in Table 6, where the outcome $S_j = 0$ is taken as a reference outcome.

Table 6: Estimation of the Multinomial Logit Model

Variable	Work and school		School only	
	Coefficient	*t*-value	Coefficient	*t*-value
Y_{-i}	0.0003	2.49	-0.0001	-0.89
w_i	-0.008	-5.77	-0.009	-12.45
total members of household	-0.101	-3.07	0.001	0.04
primary school completed	4.306	23.35	3.063	27.44
years of schooling	-0.580	-2.66	-1.474	-11.41
(years of schooling)^2	0.104	5.78	0.171	16.89
(age-years of schooling)^2	-0.033	-5.02	-0.043	-10.82
female	-0.918	-9.35	-0.647	-11.11
max parental education	0.039	1.87	0.125	10.06
children below 6 in household	-0.010	-0.15	-0.180	-4.71
rank	0.015	0.22	0.074	1.87
median state earnings	-0.561	-3.86	-0.128	-1.51

Note: Table reports estimates of the multinomial logit for the group of 12-14 years old for one residual draw. Outcome "not school" is the comparison group. 9,102 observations, Pseudo R2 = 0.11. Constant included but not reported.

On average, the model predicts correctly about 75 percent of the choices (Table 7). A breakdown by age and status reveals that the poor fit is concentrated in age/outcome cells that contain relatively few observations. This points to a potential difficulty of the multinomial logit model in dealing with unequally distributed choices.[13]

Table 7: Accuracy of Model Prediction *(K=1)*

	Predicted outcome			
Original outcome	Not school	Work and school	School only	Total
Not school	3,757	16	2,030	5,803
Work and school	152	655	1,084	1,891
School only	1,529	227	11,513	13,269
Total	5,438	898	14,627	20,963

Note: Table is a cross tabulation of the actual status and the predicted status, based on 1000 simulations. The share of correct predictions is 76 percent.

---

[13] For the youngest cohort, the estimated multinomial model would not predict any of the individuals to be in the "not school" group. Those that do not go school in this cohort are driven into this choice by a rather large individual shock rather than the structural parameters of the model.

As discussed in Section II.B, identification requires an assumption about K, which is the amount of time dedicated to household production or work in the labor market. Assuming $K = 1$, one obtains a different estimate for each age group and the structural parameters α_j and β_j. Table 8 illustrates the differences across age groups. Estimates of M range from around 0.37 to 0.46, in line with findings of Bourguignon et al. (2002). This implies that those who work and go to school realize on average around 40 percent of their potential earnings. Estimates of D are very close to unity for the youngest and eldest age groups, the low coefficient for the middle age group is somewhat surprising. The coefficients for the household income, α_j, which determine the reaction of each individual utility to income other than i's decrease significantly with age, falling from 0.21 to 0.01, but remain positive. Compared to Bourguignon et al. (2002) who also find a decreasing effect of household income with age, the results here are slightly more pronounced. Finally, the coefficients on the imputed wage have the same magnitude and range across age groups, but are more or less constant across choices but within age groups.

Table 8: Estimation of Structural Parameters (*K=1*)

Age group	M	D	Household income (Y_{-i})			Imputed wage (w_i)		
			α_0	α_1	α_2	β_0	β_1	β_2
10-11	0.432	0.988	0.211	0.211	0.211	0.211	0.091	0.208
12-13	0.372	0.292	0.013	0.013	0.013	0.013	0.005	0.004
14 -16	0.459	0.907	0.026	0.026	0.026	0.026	0.012	0.023

Note: Parameters for *K*=1, based on 1000 simulations.

These results are robust to variations of the structural parameter *K*. Table 9 depicts estimates of the structural parameters for a value of *K=0.5*, and shows that the model's prediction is not sensitive to the choice of *K*. In fact, the share of correct predictions remains unchanged. This robustness is largely due to the fact that the *difference* between the coefficients is more important than their absolute levels, as the parameter *K* affects the utility level for every choice.

Table 9: Accuracy of Model Prediction (*K=0.5*)

Original outcome	Predicted outcome			Total
	Not school	Work and school	School only	
Not school	3,753	20	2,030	5,803
Work and school	152	652	1,087	1,891
School only	1,505	223	11,541	13,269
Total	5,410	895	14,658	20,963

Note: Table is a cross tabulation of the actual status and the predicted status, for *K*=0.5 based on 1000 simulations. The share of correct predictions is 76 percent.

C. Impact Simulation

The PROGRESA transfer increases with grades, and, for the secondary school, is higher for girls than for boys. Household income support through the transfer is sizeable, with the average transfer for one child in secondary school representing 15 percent of average total household income (Table 10).

Table 10: PROGRESA Transfer Scheme

Primary school		Secondary school		
			Transfer 1/	
Grade	Transfer 1/	Grade	Boys	Girls
Three	65	One	190	200
Four	75	Two	200	220
Five	95	Three	210	240
Six	130			
Memo items				
Average monthly household income 2/				1,412
Average monthly household income per household member 2/				203

1/ Values indicate the monthly transfers in Mexican pesos payable to each eligible child on attendance of the specific grade (values valid for the first semester 1998). 10 pesos roughly corresponded to 1.10 US Dollar in that period. A maximum of 585 pesos was payable to each household each month.

2/ First wave.

For the simulation, the transfer TR_i is determined for each i by the gender and the years of schooling completed and transformed into a weekly subsidy.[14] Simulation of the transfer is straightforward. While the utility under choice $j = 0$ remains unchanged (the program's conditionality) under choices $j = 1, 2$ the term $\alpha_j TR_i$ has to be added to the index. Together with the estimated residuals this gives a new utility level $\tilde{U}_j$ and an option j is chosen such as to maximize utility. The difference of the distributions under U_j and $\tilde{U}_j$ determines the impact of the program. This is simulated for each age group and according to i's characteristics.

The impact of the program is clearly visible from Table 11, which shows the transition from the actual to the simulated status. The number of children not going to school decreases, and the decrease is captured mainly by an increase in those that go to school only, while a few individuals switch to the option work and school. For example, from the 4,661 children that did not attend school, 1,165 are expected to attend school after receiving the transfer, and 118 would attend school and work. This implies an increase in the school enrollment ratio by 7.5 percentage points. A closer look at the effect on enrollment ratios by gender and age group is deferred to Section V.

Table 11: Estimated Transition Matrix

| | Predicted outcome | | | |
Original outcome	Not school	Work and school	School only	total
Not school	3,378	118	1,165	4,661
Work and school	0	1,551	0	1,551
School only	0	9	10,816	10,825
Total	3,378	1,678	11,981	17,037

Note: Table presents the transition from the observed occupational choices (rows) to the simulated choices after receiving the transfer (columns), eligible individuals only, based on 1000 simulations.

---

[14] All household in which the transfer would have surpassed the monthly limit were dropped at the outset of the analysis (2.7 percent of households).

IV. THE BENCHMARK: EX POST ANALYSIS

PROGRESA was implemented as a randomized experiment. An initial survey conducted before the program was implemented identified eligible households and the population was divided in a treatment and a control group. While factors including the phasing-in of the program have played a role in determining the control group, the randomization is generally regarded as successful. This allows to identify the impact of PROGRESA by comparing outcomes of the treatment with the control group. The effect of the program on school enrollment ratios has been extensively analyzed in the literature. The findings of this section are in line with the previous studies, such as Schultz (2001), Behrman et al. (2001), and Attanasio et al. (2001), and will therefore be presented in a concise way. The comparison between school attendance before and after the program was implemented will be confined to making a cross section comparison between wave 5 (November 99) and waves 1 and 2.

A. Pre–Program Differences

In PROGRESA, by and large, the randomization was successful in that the control and treatment group do not exhibit large differences at the outset. A closer look at pre-program differences in school enrollment reveals that with one exception there are no significant differences. Table 12 and Table 13 report the average enrollment ratios by age in the control and in the treatment group, separately for boys and for girls. As can be seen from Table 12, while the overall sample exhibits one significant positive difference between treatment and control group in the category of 13 years old, restricting the analysis to the eligible sample none of the differences are significant. Indeed, the differences between control and treatment group are very small in value and range from -0.6 to 4 percentage points.

In contrast, for girls (Table 13) there is one significant difference in the pre-program values, where the treatment group exhibits a significantly lower value (-4.5 percent) in the category of eligible 12 years old. In the next age category (13 years) the value of the treatment group is higher (4 percent, although not significant). Except for these two cases all differences are negligible.

Table 12: Pre-Program Differences: Boys

	Age						
	10	11	12	13	14	15	16
Full sample							
Treatment	0.972	0.956	0.892	0.892	0.659	0.496	0.344
Control	0.977	0.952	0.883	0.753	0.647	0.467	0.322
Difference	-0.005	0.004	0.009	0.048	0.011	0.028	0.022
t-value	0.65	0.34	0.60	**2.33**	0.46	1.13	0.89
Not eligible							
Treatment	0.975	0.985	0.906	0.864	0.772	0.622	0.298
Control	0.975	0.961	0.924	0.773	0.712	0.540	0.253
Difference	0.000	0.024	-0.018	0.091	0.061	0.083	0.045
t-value	0.01	1.10	0.56	**2.08**	1.14	1.47	0.99
Eligible							
Treatment	0.971	0.952	0.888	0.788	0.635	0.470	0.361
Control	0.977	0.951	0.874	0.749	0.634	0.447	0.350
Difference	-0.006	0.000	0.014	0.040	0.001	0.023	0.011
t-value	0.70	0.04	0.85	1.72	0.05	0.81	0.38

Note: Table reports for each cell school enrolment ratios, the difference and the t-values. First wave only.

Table 13: Pre-Program Differences: Girls

	Age						
	10	11	12	13	14	15	16
Full sample							
Treatment	0.970	0.955	0.808	0.703	0.557	0.409	0.286
Control	0.967	0.955	0.850	0.677	0.544	0.374	0.236
Difference	0.003	0.000	-0.042	0.026	0.013	0.035	0.049
t-value	0.38	0.02	**2.20**	1.13	0.51	1.37	**2.01**
Not eligible							
Treatment	0.972	0.974	0.867	0.720	0.669	0.523	0.298
Control	0.990	0.956	0.894	0.788	0.589	0.377	0.222
Difference	-0.018	0.018	-0.026	-0.068	0.080	0.146	0.076
t-value	1.00	0.78	0.67	1.30	1.41	**2.55**	1.57
Eligible							
Treatment	0.970	0.952	0.795	0.700	0.531	0.378	0.282
Control	0.963	0.955	0.841	0.653	0.533	0.373	0.242
Difference	0.007	-0.003	-0.045	0.047	-0.002	0.005	0.040
t-value	0.75	0.25	**2.09**	1.80	0.07	0.19	1.41

Note: Table reports for each cell school enrolment ratios, the difference and the t-values. First wave only.

B. Difference Estimation

If the randomization of the experiment was successful, the difference-in-differences (DD) estimator should coincide with the simple difference estimator (D) which compares the treatment with the control group after the program completed. In particular, since both groups were affected by the same shocks and there were no significant pre-program differences between the groups, the simple D estimator is sufficient to identify the effect.

The difference estimator D compares, by age and gender, the difference between the enrollment ratios of the control and the treatment group at the end of the program. Table 14 and Table 15 provide an estimate of these differences. Table 14 shows that the enrollment ratio among eligible boys (last column) has risen, on average, by 4.9 percentage points. The increase was most pronounced in the age groups 13 to 15 years (up 8 points) and with just 2 percentage points it is lowest among the 10 to 11 years old. A look at the non-eligible population confirms that the targeting of the program was successful, because with the exception of the 13 years old these do not differ between the two populations. A similar picture emerges for the girls where the average effect was higher with 6.1 percentage points. Among the eligible the effect was highest for the 13 to 15 years old and, as for the boys, lowest in the youngest age group, where it is also insignificant.

The D and DD estimator produce very similar results, especially for girls (Figure 1). Notice also that the small kink in the DD estimate for 13 years old is a reflection of the pre-program differences in that particular age group. For boys, the D estimate is slightly above the DD estimate but does follow the same pattern.

Table 14: Difference Estimator: Boys

|				Age				
	10	11	12	13	14	15	16	10-16
				Full sample				
Difference	0.017	0.019	0.050	0.045	0.040	0.082	0.041	0.043
t-value	**2.52**	**2.36**	**3.69**	**2.44**	1.79	**3.18**	1.60	**5.46**
				Not eligible				
Difference	-0.003	0.010	0.041	-0.094	-0.041	0.065	0.031	0.009
t-value	-0.19	0.44	0.96	**-2.00**	-0.71	1.08	0.50	0.43
				Eligible				
Difference	0.020	0.020	0.051	0.070	0.054	0.085	0.044	0.049
t-value	**2.69**	**2.37**	**3.60**	**3.47**	**2.25**	**2.99**	1.55	**5.78**

Note: Table reports for each cell the difference between in enrolment ratios between the treatment and the control group respectively, and the corresponding t-values.

Table 15: Difference Estimator: Girls

				Age				
	10	11	12	13	14	15	16	10-16
				Full sample				
Difference	0.001	0.006	0.061	0.084	0.077	0.069	0.058	0.053
t-value	0.14	0.62	**3.90**	**4.16**	**3.20**	**2.64**	**2.11**	**6.26**
				Not eligible				
Difference	0.015	-0.032	0.079	0.105	-0.042	-0.048	0.041	0.008
t-value	0.75	-1.38	1.81	**1.93**	-0.74	-0.81	0.65	0.37
				Eligible				
Difference	-0.001	0.013	0.059	0.081	0.102	0.098	0.062	0.061
t-value	-0.21	1.15	**3.48**	**3.70**	**3.84**	**3.37**	**2.04**	**6.69**

Note: Table reports for each cell the difference between in enrolment ratios between the treatment and the control group respectively, and the corresponding t-values.

Figure 1: Enrollment Ratio: Actual Effect of PROGRESA, D and DD Estimator
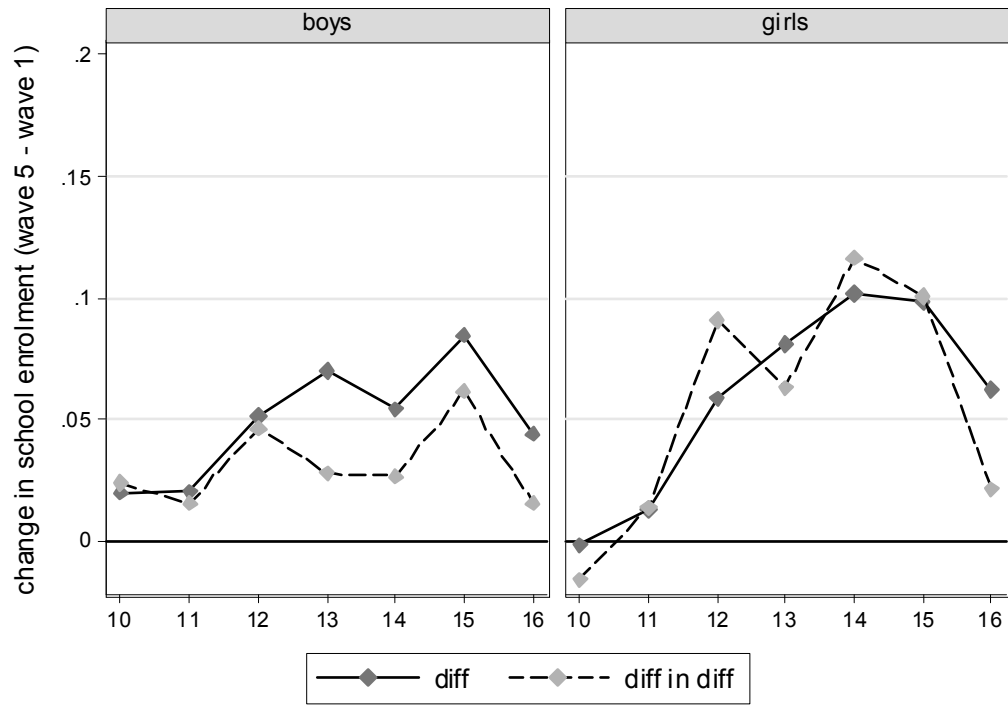
V. COMPARISON OF RESULTS AND DISCUSSION

This section compares the ex ante and the ex post evaluation results. For this purpose the trichotomous choice variable is reduced by the work dimension so that it only reflects whether a child attends school or not. The results of the simulation, the difference and difference in difference estimator, by gender and age, are presented in Table 16. Figure 2 shows this graphically, including 90 percent confidence intervals.[15]

In general, the simulation is close to the actual effect, and can be regarded as a good approximation of the actual effect of the program. In particular:

- While the difference estimator puts the increase in overall school enrollment ratio at 5.4 percentage points, the simulated effect is estimated to be somewhat higher at 8.1 percentage points. The deviations seem to be highest for the age group of 15 and 16 years. However, in 9 of the 14 gender/age subgroups the actual effect is within a 90 percent confidence interval of the simulation.

- The simulation results are close to the actual for middle age group (ages 12-14), the differences for the young age group are also small, but significant. The deviations between simulation an actual effect are largest for the eldest group (ages 15 and 16).

- The simulated and the real effect follow a similar pattern across ages and gender; the model seems to capture the underlying age-gender differences well.

- The estimation does not capture the transition to secondary school well. Differences between the simulation and actual effect are largest among those who have just completed primary school (Table 17). For example, the enrollment ratio for 16 years olds that have completed primary school increased from 9 to 12 percent during PROGRESA, whereas the model would predict an increase to 25 percent.

- The simulation is robust to various choices of K. Simulations for smaller K lead to smaller effects (Table 16), but the impact of K is small relative to the difference between the overall simulation and the actual effect.

---

[15] Confidence intervals were obtained using the bootstrap mechanism outlined in Appendix D.

Figure 2: Simulated effect and D estimate
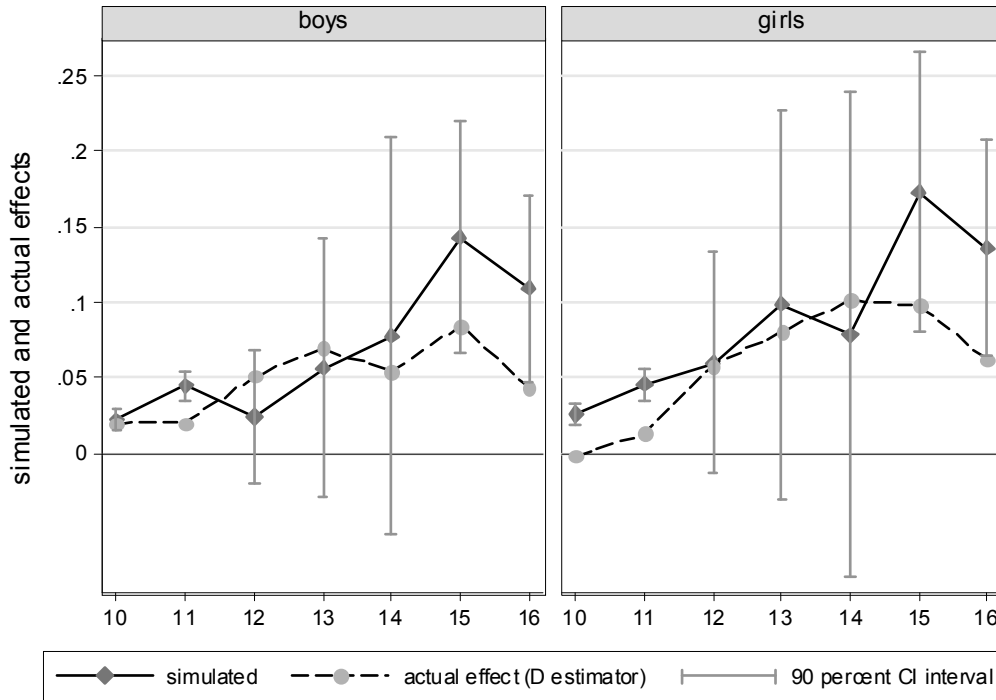
- The estimated transition matrix (Table 11) shows that the program will also reduce the incidence of child labor. Some 90 percent of the children that are predicted to resume schooling are also predicted not to participate in the labor market, whereas 10 percent will do both, go to school and work. [16]

- The poor predictive power of the multinomial choice model for some age/choice combinations negatively affects the overall result of the simulation. As can be seen from Table 2, the three outcomes of the dependent variable are quite unequally distributed for certain age groups: for example, among the 10 and 11 years old, school attendance is clearly dominating and the fraction of children that work but do not go to school is very small. Among the 15 to 16 years old, the fraction of those who "work and go to school" is very small.

---

[16] Specifically, of the (1,165+118) children that will attend school, 1,165 will not work.

Table 16: Simulation, D and DD Estimates, and Sensitivity to *K*

		Age							
		10	11	12	13	14	15	16	10-16
Boys	Before	0.974	0.953	0.885	0.788	0.646	0.476	0.361	0.726
	Simulated impact	0.022	0.044	0.033	0.054	0.062	0.158	0.118	0.070
	D impact	0.02	0.02	0.051	0.07	0.054	0.085	0.044	0.049
	DD impact	0.024	0.016	0.046	0.028	0.026	0.062	0.015	0.031
Girls	Before	0.967	0.953	0.815	0.687	0.538	0.389	0.271	0.660
	Simulated impact	0.027	0.046	0.061	0.087	0.103	0.191	0.127	0.092
	D impact	-0.001	0.013	0.059	0.081	0.102	0.098	0.062	0.059
	DD impact	-0.016	0.014	0.091	0.063	0.116	0.101	0.021	0.056
Total	Before	0.971	0.953	0.85	0.738	0.592	0.433	0.316	0.693
	Simulated impact	0.024	0.045	0.047	0.071	0.083	0.174	0.123	0.081
	D impact	0.009	0.017	0.055	0.075	0.078	0.091	0.053	0.054
	DD impact	0.004	0.015	0.068	0.046	0.071	0.081	0.018	0.043
Total	Simulated impact (*K*=1)	0.024	0.045	0.047	0.071	0.083	0.174	0.123	0.081
	Simulated impact (*K*=0.75)	0.025	0.044	0.046	0.075	0.08	0.168	0.124	0.080
	Simulated impact (*K*=0.5)	0.024	0.045	0.042	0.078	0.078	0.158	0.123	0.078

Note: Table reports the enrolment ratios at the beginnning and the impact according to the simulation (1000 simulations), the D, and DD estimator.

Table 17: Transition to Secondary School

Years of schooling completed			Age					
			12	13	14	15	16	10-16
5		Enrollment ratio (wave 1)	94.9	91.1	79.1	52.3	22.9	89.0
		Enrollment ratio (wave 5)	96.9	93.6	80.3	57.9	31.8	92.8
		Enrollment ratio (simulated)	96.1	93.1	85.4	69.4	42.0	92.6
6	(Primary school completed)	Enrollment ratio (wave 1)	73.1	51.7	29.7	17.8	9.3	39.3
		Enrollment ratio (wave 5)	81.8	64.0	38.6	21.4	11.9	50.7
		Enrollment ratio (simulated)	81.1	64.7	44.8	46.8	25.5	55.6
7		Enrollment ratio (wave 1)	98.5	97.1	94.6	91.2	78.2	94.0
		Enrollment ratio (wave 5)	96.5	97.9	96.0	84.5	74.7	93.0
		Enrollment ratio (simulated)	99.0	98.0	96.5	94.7	87.2	96.3
8		Enrollment ratio (wave 1)	.	96.1	96.1	92.0	87.2	93.3
		Enrollment ratio (wave 5)	.	98.8	96.9	90.9	86.8	93.9
		Enrollment ratio (simulated)	.	97.2	98.1	95.3	91.7	96.0

Note: Table is a cross tabulation of the age and education, indicating for each cell the actual and simulated enrollment ratios.

VI. CONCLUDING REMARKS

This paper applies a microsimulation method to evaluate the impact of a conditional cash transfer program and compares the effect with the actual outcome of the policy. The simulation correctly predicts that school enrollment ratios among the target population will increase as a result of a cash payment for school attendance. Hence, the model is a valuable tool to asses the effects of such schemes which, by increasing educational attainment, are also likely to reduce poverty in the long run. A discrete model of occupational choice is used which allows for three outcomes: attend school, work, or do both. The model picks up the main mechanism through which the transfer affects the schooling decision, considering explicitly actual and potential market earnings of each child. While the overall performance of the microsimulation is good, two potential shortcomings stand out. First, the simulation does not perform well if the population is unequally distributed over the three categories. This is a common problem of choice models. One solution—and possible extension of this paper —is to employ a microsimulation method for a dichotomous framework. While this comes at the cost of loosing important features of the estimates such as the effect of the program on the amount worked, it might improve the forecasting abilities with respect to school enrollment ratios. Second, the model does not capture well the transition to secondary school, pointing to another area of improvement.

VII. REFERENCES

Anker, R. (2000): *Conceptual and Research Frameworks for the Economics of Child Labor and its Elimination*, IPEC Working Paper, Geneva: ILO.

Attanasio, Orazio; Meghir, Costas and Santiago, Ana (2001): *Education Choices in Mexico: Using a Structural Model and a Randomized Experiment to Evaluate PROGRESA*, Mimeo.

Behrman, Jere; Sengupta, Piyale and Todd, Petra (2001): *Progressing Through PROGRESA: An Impact Assessment of a School Subsidy Experiment*, mimeo.

Bourguignon, François; Ferreira, Francisco H. G. and Leite, Phillippe (2002): *Ex ante Evaluation of Conditional Cash Transfer Programs: The Case of Bolsa Escola*, World Bank Policy Research Working Paper Nr. 2916, Washington, DC: The World Bank.

Bourguignon, François; Fournier, Martin and Gurgand, Marc (2001a): *Selection Bias Correction Based on the Multinomial Logit Model*, mimeo.

Bourguignon, François; Fournier, Martin and Gurgand, Marc (2001b): "Fast Development with a Stable Income Distribution: Taiwan, 1979 - 1994", *Review of Income and Wealth*, 47 (2): 139–163.

Heckman, J. (1979): "Sample Selection Bias as a Specification Error", *Econometrica*, 47: 153–161.

ILO (2003): *Investing in Every Child: An Economic Study of the Costs and Benefits of Eliminating Child Labor*, Geneva: International Labor Organization.

Lalonde, Robert (1986): "Evaluating the Econometric Evaluations of Training Programs with Experimental Data", *American Economic Review*, 76 (4): 604–620.

Lee, L. F. (1983): "Generalized Econometric Models with Selectivity", *Econometrica*, 51: 507–512.

McFadden, Daniel (1973): "Conditional Logit Analysis of Qualitative Choice Behavior ", in: "Frontiers in Econometrics", (ed. ZAREMBKA, P.), New York: Academic Press.

OECD (2003): *Combating Child Labor: A Review of Policies*, Paris: Organisation for Economic Cooperation and Development.

Schultz, Paul T. (2001): *School Subsidies for the Poor: Evaluating the Mexican Progresa Poverty Program*, Economic Growth Center Discussion Paper Nr. 834, New Haven: Yale University.

Todd, Petra and Wolpin, Kenneth I. (2003): *Using a Social Experiment to Validate a Dynamic Behavioral Model of Child Schooling and Fertility: Assessing the Impact of a School Subsidy Program in Mexico*, mimeo.

Worldbank (2000): *World Development Report: Attacking Poverty*, Washington DC: World Bank.

VIII. APPENDIX

A. Key Elements of PROGRESA

PROGRESA is the acronym of *Programa de Educación, Salud y Alimentación* and is a program that aims at developing the human capital of people living in poor rural households in Mexico.[17] Launched by the federal government of Mexico in 1997, the International Food and Policy Research Institute (IFPRI) joined the effort a year later together with other research and development institutions. The program is still running band is now known as OPORTUNIDADES. One of the main objectives of the program is to improve the school attendance of children. Eligible households with school aged children receive grants conditional on school attendance. The size of the grant increases with the grade (starting from the third year of primary school) and, for secondary education, is slightly higher for girls than for boys (see Table 10). Initially, 506 localities were chosen to participate in the program. For logistic reasons and evaluation purposes, the sample was divided into to a treatment (320) and a control (186) group, where the program started two years later. The selection into the treatment and control group can be considered as having been random - at least with respect to the variables that interest in this analysis. Within each village the survey covers all households (roughly 24,000 observations) and collects extensive information on consumption, income, nutrition and other issues. For each *household member*, including each child, there is information about age, gender, education, labor supply, income (various forms), school enrollment, nutrition, and health status. Detailed information about the localities is also available. However, the questionnaire used varies substantially between waves. Based on the information collected in the first round of interviews an eligibility criterion was established and the sample was classified into eligible and non eligible households. Later the eligibility was extended (known as the *densificación*) such that now finally some 80 percent of the households were eligible to participate in the program.

---

[17] For further details and documentation see www.ifpri.org/themes/progresa.htm.

B. Data Description

The data of the first five waves is publicly available from the website of IFPRI.[18] The following is a summary of how the variables used in the analysis were constructed from the survey questions in wave 1. Expressions such as *p08* refer to the number in the questionnaire and the variable name in the original data set.

S_i is the occupational choice variable. It takes value 0 if the answer to *p21* is "No", it takes value 1 if the answer to *p21* is "Yes" and any of the answers to *p22*, *p23*, *p301* or *p302* indicate that *i* works for income. It takes value 2 if the answer to *p21* is "Yes" and the other aforementioned questions do not indicate labor income. If someone does not report income but reports to be working observations are dropped.

Y_i is the weekly monetary income of *i*'s household, i.e. $\sum_{i=1}^{N} y_i$

y_i denotes the income of individual *i*. Using all sources of monetary income (questions *p22*, *p23*, *p301* and *p302*), the payment received (*p291m*, *p31a2*, *p31b2*), the period of payment (*p291p*, *p31a1*, *p31b1*) and the amount of hours worked (*p2612*) a weekly income variable for each individual *i* is constructed.

Y_{-i} is the household's income without *i*'s contribution, i.e. $Y_i - y_i$.

age_i refers to question *p08*.

$female_i$ refers to question *p11*.

edu_i refers to question *p20*. Years of schooling completed are calculated from *primaria* onwards, where each of the 6 *niveles* counts as one year. Hence after completion of the *secundaria* (3 *niveles*), one has 9 years of schooling.

---

[18] See http://www.ifpri.org/data/dataset.htm.

rank$_i$ is computed as the number of household members that are older than i but less than 19 years old plus one.

N$_i$ is the total number of persons living in the household.

child$_i^{<6}$ is the total number of children below the age of 6 in the household of i.

poor$_i$ indicates if i belongs to an eligible household, using *pobreden*.

treat$_i$ indicates if i belongs to a treatment village, using *contbas2*.

state_w$_i$ the median earnings of i's gender/age group in his or her respective state. There are 7 states in the sample.

For the ex post analysis, in addition to the aforementioned variables from wave 1, the analogous variables from waves 2 and 5 for age, schooling, gender etc. were used.

C. Method for Drawing Choice-consistent Residuals

For notational convenience, drop index i and call $x\beta$ the probability index that captures all variables. Then the underlying random utility model takes the form

$$U_j = x\beta_j + u_j$$

Assuming that u_j derives from a double exponential distribution with independent draws, the choice probabilities p_j can be written as:[19]

$$p_j = \mathrm{Prob}(x\beta_j + u_j > x\beta_k + u_k \text{ for all } k \neq j)$$

$$p_j = \frac{\exp^{x\beta_j}}{\sum_{k=o}^{2} \exp x\beta_\partial}$$

The distribution of u_j conditional on the a choice $k \neq j$ has the following c.d.f.:

$$F(u_j \mid S_j = k) = \frac{\exp^{-\exp^{-u_j}}}{\exp^{-\exp^{x\beta_j - x\beta_k - u_k}}}.$$

The inverse of this function is used to draw residuals in the following sequence:

$$u_k = -\log(-p_k \cdot \log(rnd())) \text{ if } j = k$$

$$u_j = -\log(\exp^{-u_k} \cdot (p_j / p_k)) - \log(rnd()) \text{ if } j \neq k$$

Where $rnd()$ is a random number between $[0,1]$.

D. Bootstrap Mechanism

The algorithm consists of 4 steps.

Step 1 Compute the original sample estimates: estimation of equations (4) and (3) and simulation of the program.

Step 2 Random draw with replacement of a new sample containing as many observations as the original sample and compute all estimates described in step 1.

Step 3 Repeat step 2 for 1000 times collecting the predicted values of the age/gender specific impact. Drop the upper and lower five percentile of this distribution.

Step 4 Use the distribution of values obtained in step 3 to obtain confidence intervals around the estimates obtained in step 1.