



How do *Ex Ante* Simulations Compare with *Ex Post* Evaluations? Evidence from the Impact of Conditional Cash Transfer Programs*

Phillippe LEITE, Ambar NARAYAN & Emmanuel SKOUFIAS**

Abstract: We compare the *ex ante* simulation of the impacts of conditional cash transfer (CCT) programs against the *ex post* estimates of impacts obtained from experimental evaluations. Using the data on program-eligible households in treatment areas from the same baseline surveys used for the experimental evaluations of CCT programs in Mexico and in Ecuador, we use a micro-simulation model to derive *ex ante* estimates of the impact of the programs on enrollment rates and poverty. Our estimates reveal that *ex ante* predictions of certain impacts of CCT programs using the micro-simulation model match up well against the benchmark estimates of *ex post* experimental studies. The findings seem to support the use of this model to assess the potential impact and cost efficiency of a CCT program *ex ante*, for the purpose of informing decisions about the design of the program.

Keywords: Conditional Cash Transfers, Micro-simulation, Experimental evaluation

* This is an open-access article distributed under the terms of the Creative Commons license Attribution-Noncommercial-NoDerivs 3.0 Unported. Distribution and reproduction are permitted, provided that the original authors and JPAID as the source are credited. The authors are grateful to Francisco Ferreira for his inputs in the early stages of this work, to Tahrat Shahid for excellent research assistance, and to the Multi-Donor Trust Fund for Poverty & Social Impact Analysis for partial support.

** The authors share equal responsibility for the contents of this paper. The findings, interpretations, and conclusions are entirely those of the authors. They do not necessarily represent the view of the World Bank, its Executive Directors, or the countries they represent. Emmanuel Skoufias is corresponding author. The World Bank, MSN 14-405, 1818 H Str. NW, Washington D.C., 20433. Email: eskoufias@worldbank.org. Tel: (202) 458-7539.

Manuscript received 01.09.2014; revised 01.06.2015; accepted 10.06.2015

jpaid.yonsei.ac.kr

Conditional Cash Transfer programs (CCTs) have been used by many governments in their attempts to reduce poverty and inequality and achieve longer term human development. While CCT programs in different countries may vary in their design and purpose, they tend to share a few basic features. All involve transferring resources (usually in the form of cash) to poor households, which are usually provided directly to the mother with reciprocal conditions imposed on the household, to encourage changes in behavioral norms. These conditions commonly include requiring, for example, that children attend school and/or individuals have regular health facility visits.

Different methods of evaluation have been developed to analyze the likely impacts of a monetary transfer on human capital accumulation and poverty alleviation. The array of literature on program evaluation presents two broad types of analysis, *ex post* and *ex ante*. *Ex post* evaluations can in turn be divided into two categories: experimental and quasi-experimental. Experimental analyses are those where households or geographic units are randomly assigned to both treatment and control groups prior to the implementation of the program. Quasi-experimental evaluations are those where treatment and control groups are *not* selected through random assignment (of households or geographic units), with the result that statistical techniques are necessary to correct and control for observable and unobservable differences between the two groups that might affect outcomes of interest.

Ex post evaluation methods have major advantages in terms of generating statistically valid results about the impact of a monetary transfer through a CCT program, against a counterfactual of not having the program. However, such methods can also have a few important disadvantages, depending on the type of techniques used. Experimental *ex post* evaluations may provide more statistically valid and robust results relative to quasi-experimental methods, but they are often costly and time-consuming to implement, particularly on a large scale. Quasi-experimental techniques, while typically less costly and easier to implement than experimental analyses, have the disadvantage of

requiring a high level of statistical complexity to account for selection bias and the effects of unobservable characteristics.

Regardless of the method used, *ex post* evaluations by themselves are also likely to provide little evidence prior to implementation about the *potential* impact of programs, which is what policymakers designing large programs or reforms are most interested in. Often the design and implementation of large income transfer programs cannot be delayed until a series of pilots “testing” various design elements have been completed and evaluated using *ex post* techniques. In other cases, even when pilots have been tested and evaluated, concerns about “external validity” imply that the policymaker cannot be sure whether the results from a pilot will be valid for a program implemented on a much larger scale. Moreover, policymakers are often concerned about sensitivity of outcomes to transfer amounts or the relative merits of alternative program designs. While testing alternative program designs is possible in theory using *ex post* methods, implementing and evaluating the relative impacts of multiple “treatments” tends to be extremely difficult in practice.

Ex ante evaluations involve simulating the effects of a program on the basis of a household model, in most cases using a data set representative of all program beneficiaries. Such techniques, while having the disadvantage of not measuring the actual impacts of a program, can be particularly useful in addressing some of the questions mentioned above. At the same time these techniques, which rely on structural models of economic behavior, are often criticized for their reliance on strong underlying assumptions. Bourguignon, Ferreira and Leite (2003) (henceforth BFL) propose such a behavioral model that allows policymakers to evaluate the potential effect of a CCT on poverty, inequality and school enrollment rates. The model combines arithmetic methods with models of demand for schooling, generating predictions about the potential occupational choice (one of which is the decision of going to school) of a child living in a region where a CCT is implemented. The BFL model generates estimates of likely effects of a

monetary transfer program, comparing the simulated occupational choice of children with the status quo (absence of the CCT), keeping everything else constant.

Ravallion (2008) underscores that rather than being competing methods, *ex ante* and *ex post* methods can usefully complement each other. For example, combining an *ex post* evaluation with an *ex ante* structural model of schooling choices would allow policymakers to expand the existing set of policy alternatives in considering the optimal design of a given program. That said, for the two approaches to complement each other in the context of a specific program, consistency of results must be an important consideration. The concerns about assumptions that underlie the *ex ante* behavioral models make it all the more important that their predictions are checked for consistency with the results of *ex post* experimental evaluations. Any evidence to suggest that *ex ante* approaches yield results that are at least broadly consistent with those from methodologically sound experimental evaluations would strengthen the case for using such approaches to complement *ex post* evaluations of similar programs, where the relative ease of employing *ex ante* approaches can be particularly useful in evaluating alternative design options before a program is implemented.

Our paper is an attempt to generate evidence on the consistency between the two approaches in the context of CCT programs by comparing *ex ante* predictions with results of *ex post* experimental evaluations, following the suggestion of Bourguignon and Ferreira (2003). Specifically, we compare results from *ex ante* simulations of the likely impact of *PROGRESA* in Mexico and *Bono de Desarrollo Humano* (henceforth BDH) in Ecuador on enrolment rate and poverty with results from well-regarded *ex post* evaluations of these programs. The *ex ante* simulations are conducted on the same baseline surveys that are used for the experimental evaluations to provide a consistent benchmark for the comparison.

The first section below reviews the literature most relevant to our analysis. The second section provides a synopsis of the BFL model,

including its underlying assumptions. The third section discusses the two CCT programs on which the BFL model will be applied (*PROGRESA* and BDH), the data to be used for the analysis and the *ex post* evaluation studies that provide benchmark results. The fourth section presents the results of the comparison between the predictions of the BFL model and results of *ex post* evaluations. The final section concludes the paper.

Review of the Literature

Cash transfer programs are noncontributory cash grants to selected beneficiaries, in the sense that beneficiaries do not pay into a system that later awards them the transfers.¹ This definition includes noncontributory pensions, poverty-based transfers, and family grants. Beneficiaries can work to receive transfers from public works schemes; and in the case of CCTs, beneficiaries might have to fulfil other requirements such as school attendance or regular health check-ups for children in order to be eligible for transfers. Most poverty-based transfers seek to target the poor (and/or specific vulnerable groups) using a suitable method for identifying the beneficiaries.

Since the early 2000's, the number of countries implementing cash transfer programs has increased exponentially and these have become the centerpiece of social protection systems of many middle-income countries. Alderman and Yemtsov (2014) show the different pathways through which cash transfers can improve household resilience and contribute to poverty reduction and growth: enabling households to make better physical and human capital investments; allowing households to diversify into higher risk, higher return activities; and promoting the creation of community assets through public works.

Since we use results from experimental *ex post* evaluations as the

¹ See Grosh et al. (2008), Garcia and Moore (2012) and Fiszbein et al. (2009).

yardstick to test the validity of *ex ante* predictions, it is useful to review some of the seminal literature concerned with the predictive power of non-experimental estimators against the benchmark of experimental evaluation results.

The act of forecasting outcomes is frequently subject to criticism, as it relies on estimation of parameters that are sensitive to model specification and to non-observable errors. Forecasts about the impacts of social programs are typically validated by comparing results with the estimated treatment effects from randomized experiments when they are available. Examples of such analyses include Todd and Wolpin (2006), Attanasio et al. (2005), and de Janvry and Sadoulet (2006) for the Mexican case, and Lise et al. (2005) for the Canadian Self-Sufficiency Project (SSP).

Todd and Wolpin (2006) use an *ex ante* model to generate child schooling estimates that compare reasonably well with experimental results. For children aged 12–15 years, the model predicts the effects of *PROGRESA* on school enrollments to be in the same direction as experimental results, but underestimates them by 21–25 percent. However, when the sample is disaggregated into two age groups, their *ex ante* model does not predict any change in outcomes for boys aged 12–13 years and significantly overestimates the effects of the program for the 14–15 years age group. Todd and Wolpin also use a particular sample that is not comparable with the one used by Skoufias and Parker (2001) and Schultz (2000) to generate *ex post* results. They restrict their sample from the original 24,077 households surveyed under *PROGRESA* to only landless households where the spouse of the household head is a woman under 50 years of age, which reduces their sample to around 3,200 households.

Attanasio et al. (2005) present a complex model to simulate *ex ante* the likely effects of *PROGRESA*, as well as that of a hypothetical program that differs from *PROGRESA* in the way the transfers vary by the child's school year. They do not compare the results of the first exercise with *ex post* results and instead only present the *ex ante*

program impacts for different age groups. According to our computations using their estimates, the change in enrollment due to *PROGRESA* predicted by their model translates to enrollment increases of approximately 2 and 5.4 percent among children of age 10–13 years and 14–17 years respectively.² In comparison, the enrollment increases attributed to *PROGRESA*—as estimated by the *ex post* evaluations of Skoufias and Parker (2001) and Schultz (2000)—are 2.4 and 7.5 percent for children of age 10–13 years and 14–17 years, respectively. The Attanasio et al. model is thus able to generate predictions that match up quite well with *ex post* evaluation results, even as its complexity and challenges of implementation are considerable.

Leite (2007) applies the BFL model to data from PNAD (*Pesquisa Nacional por Amostra de Domicílios*) 1999 in Brazil to forecast an increase of 3.9 percent in the enrollment rate of poor children of age 10–15 years. The forecast from the BFL model is close to the results from a quasi-experimental evaluation of the Bolsa Escola program by Cardoso and Souza (2004), in which a matching estimator suggests a 3.1 percent increase in the enrollment rate in boys and 3 percent increase for girls aged 10–15 years associated with the Bolsa Escola. De Janvry and Sadoulet (2006) also use *PROGRESA* data to develop a predictive model. They analyze the potential for increasing school attendance in order to help increase the efficiency of such transfers, concentrating on secondary school attendance and the extent to which it can be affected by a conditional transfer.

Lise et al. (2005) use data from the Canadian SSP to construct a dynamic, partial equilibrium model to simulate the effects of the program on labor market behavior and compare these to observed treatment effects on individuals. The SSP is a social program providing time-limited earnings supplements to income assistance recipients who

² Our computations involve extrapolating from the enrollment difference presented in Figure 1 of Attanasio et al. (2005), by assigning weights to each age group in Figure 1 and using the baseline survey.

are able to obtain full time employment within a 12 month period. Recipients of SSP were picked at random from among those eligible for these benefits. Lise et al. restrict their sample to single mothers only, with around 2,300 women in the control and treatment groups alike. Their *ex ante* model is calibrated to control group data and the experiment is simulated within this model to imitate the “welfare-to-work transition of the treatment group.” The results are then compared with those of an *ex post* evaluation of the SSP, which produces similar results.

An earlier body of literature, including LaLonde (1986), Heckman and Hotz (1989) and others, use experimental results as benchmarks for evaluating the predictive power of partial equilibrium non-experimental estimators, mainly in the context of labor market programs. These papers concluded that non-experimental methods were not effective in evaluating program impacts. In more recent literature, propensity score matching is extended with kernel and local linear matching estimators, which use multiple nonparticipants to estimate the outcomes of the control group as opposed to pair-wise estimation. This method, used by Heckman et al. (1997; 1998), was implemented with longitudinal and cross-sectional data from the National Supported Work Demonstration, another labor market program.³ The results suggest that such estimators are able to replicate experimental results only when the data examined are very similar—from the same source(s), with treatment and control groups from the same geographic labor markets, and incorporating a range of variables that influence both participation and labor market outcomes.⁴

McKenzie et al. (2006) build on the findings of previous authors: the best non-experimental estimates are obtained when the treatment and control groups are drawn from the same labor markets, data sources

³ NSW was a subsidized work experience program providing recipients with training and assistance in finding regular jobs.

⁴ Summarized from Smith and Todd (2005).

used are the same, and groups' characteristics are similar in that the "likelihood of receiving the treatment is similar in both groups." They compare experimental and non-experimental methods in analyzing the income gains from migration, using a survey on the quota of Tongans allowed to immigrate to New Zealand each year on the basis of a lottery from an excess number of applicants. Because migrants are more likely to have certain observable *and* unobservable characteristics (such as ability and drive), the non-experimental methods are shown to overstate income gains from migration by between 9 and 82 percent.

The *ex ante* labor market models suggest the difficulties in predicting the impacts of a program using a model representing standard economic incentives incorporated through wages. In our paper, however, the effects of programs are simulated using a model incorporating *household* income, where the incentives operate through channels that are quite different.

Against the backdrop of the existing literature discussed above, the main contribution of our paper is in testing the validity of an *ex ante* tool for evaluating the impacts of CCT programs (the BFL model) that is operationally useful, both in terms of ease of implementation and data requirements. Compared to the dynamic *ex ante* models suggested by Attanasio et al. and Todd and Wolpin (see above), the BFL model makes fewer demands on data requirements and does not rely on the availability of panel data, as it employs a reduced-form approach that imposes simplifying assumptions on household behavior. If the predictions of the BFL model with all its simplicity were to compare well with the impacts measured by rigorous *ex post* evaluations, it would be a practical and credible *ex ante* tool for policymakers to employ during the design and planning phase of CCT programs.

***Ex Ante* Simulation with the BFL Model⁵**

This section presents a summary description of the BFL model, used as the *ex ante* evaluation tool in this paper (see the original BFL paper for details). Rather than constructing a complete structural model of demand for schooling and intra-household labor allocation, the BFL model aims to obtain reasonable orders of magnitude for the likely effects of cash transfers that are conditional on school attendance.⁶ The structural aspects of modeling are kept to the minimum necessary to capture the main effects of the program. There are four key underlying assumptions. Firstly, occupational choice is assumed to reflect (in reduced form) the end result of whatever decision-making process about a child's time allocation unfolds in a household, regardless of *how* the decision is made. Secondly, the decision to send a child to school is assumed to be made *after* all occupational decisions by adults within the household have been made, so that the schooling decision does not affect those earlier decisions. Thirdly, the model does not allow for decisions about the occupational choice of multiple siblings in a household to be made simultaneously or jointly. Fourthly, the composition of the household is taken as exogenous. While each of these assumptions is an abstraction from reality (at least for some

⁵ A toolkit for running the BFL model was prepared by the World Bank. The toolkit provides a how-to guide in the application of the BFL micro-simulation model that enables users to analyze and compare several alternative scenarios for program design, using representative household survey data. The toolkit can be downloaded at www.worldbank.org/safetynets under toolkit.

⁶ While there is a large literature on demand for schooling, the empirical analysis of parental motives is always difficult because fully structural models to test such behavior are difficult to estimate. BFL adopts a simpler and operational approach, which does not attempt to take into account parental motives (such as the intrinsic value that they might attach to a child's education), beyond what is implicitly revealed by the occupational choice of children in the reduced form specification.

households), they make for a simple, reduced-form approach that is tractable with the baseline data available for most programs.

Let $j = 0$ be the occupational category of “not attending school,” $j = 1$ be that of “attending school and working,” and $j = 2$ be that of “attending school only.” Following the approach adopted in BFL, the utility function of child i corresponding to each occupational category j is given by the following:⁷

$$U_i(j) = Z_i \cdot \gamma_j + (Y_{-i} + y_{ij}) \cdot \alpha_j + v_{ij} \quad (1)$$

Where Z_i are the characteristics of both the child and the household, Y_{-i} is the household income without the child’s earnings, y_{ij} is the child’s income earned in alternative j ; and v_{ij} is the random variable representing idiosyncratic preferences. γ_j and α_j are parameters specific to occupational category j . The model is parsimonious in its representation of occupational choice of children and allows for all possible trade-offs between current income of the household and the schooling of the child.⁸ A key variable is the child’s earning y_{ij} since transfers depend on household income and the schooling/occupational status of a child, which in itself affects the child’s earning.

The model implicitly treats the child’s number of hours of work as a discrete choice between three alternatives, corresponding to the three occupational categories given by j . It seems reasonable to define child i ’s earning in occupational category 0 as equal to her observed market earnings denoted by w_i , since no time is spent on schooling. Assuming

⁷ See equation 2 in the BFL paper, which in turn is the linearized version of a more general specification given by equation 1 in their paper.

⁸ Note that the model can also implicitly represent a trade-off between current income of the household and future income of the child (and perhaps the household as well), since schooling would be expected to raise his or her future income.

that earnings can be characterized by the standard Becker-Mincer human capital model, w_i is given by:

$$\text{Log } w_i = X_i \cdot \delta + m \cdot \text{Ind}(j = 1) + u_i \quad (2)$$

Where X_i is a set of individual characteristics (including age and schooling achieved) and u_i is a random term that stands for unobserved determinants of earnings. The second term on the right hand side is a dummy variable representing the fact that a child who attends school and works for wages has less time available and is therefore likely to earn less. The child's contribution to the household income (y_{ij}) in the various alternative categories for j is given by:

$$y_{i0} = w_i; \quad y_{i1} = M \cdot y_{i0} = M w_i; \quad y_{i2} = D \cdot y_{i0} = D \cdot w_i; \quad (3)$$

$$\text{with } M = \exp(m)$$

Where y_{ij} covers income from both market-based and domestic work done by child i in occupational category j . Being in category 1, namely attending school while working outside the household, leads to a reduction in total income (relative to income in category 0) by the proportion $(1-M)$. Similarly, going to school without working in the market leads to a reduction in total income by the proportion $(1-D)$.⁹ While D is not observed, M is assumed to be the same for domestic and market work and can be estimated from observed earnings represented by equation 2 above. Replacing 3 in equation 1:

$$U_i(j) = Z_i \cdot \gamma_j + Y_{-i} \cdot \alpha_j + w_i \cdot \beta_j + v_{ij} \quad (4)$$

$$\text{where } \beta_0 = \alpha_0; \quad \beta_1 = \alpha_1 \cdot M \text{ and } \beta_2 = \alpha_2 \cdot D$$

⁹ Note that total income in category 1 can be a combination of incomes from market and domestic work of the child, while total income in category 2 is from domestic child work only.

Child i (or the household of child i) will choose the occupational category j that yields the highest utility among the three alternatives. Assuming that all parameters of equation 2 are known, along with actual or potential earnings (w_i) and the residual terms (v_{ij}), household's selection of a child's occupation can then be defined by:

$$k^* = \text{Arg max } [U_i(j)] \quad \forall j = 0, 1, 2 \quad (5)$$

If a CCT program is implemented where *all* children going to school receive a transfer T , equation 4 becomes:

$$U_i(j) = Z_i \cdot \gamma_j + (Y_{-i} + T_{ij}) \cdot \alpha_j + w_i \cdot \beta_j + v_{ij} \quad (6)$$

$$\text{where } T_{i0} = 0 \text{ and } T_{i1} = T_{i2} = T$$

Equation 6 adds a transfer amount T to the part of household income that is independent of the child's income, conditional on the child going to school. Thus in maximizing the utility function given by equation 6, the household must take into account the fact that it will receive T only in states ($j = 1$) or ($j = 2$). Under the assumptions made earlier, equation 6 is the full reduced-form model of the occupational choice of children, which would allow for simulations of the impact of CCT transfers on those choices, once the estimates of the parameters of the model as well those of w_i and v_{ij} 's are obtained.

Estimation of discrete choice model

An assumption that the v_{ij} are independently and identically distributed across sample observations with a double exponential distribution leads to the well-known multinomial logit model, which can be used to estimate equation 6, where $j = 0, 1, 2$ are the three alternative categories of occupation the household can choose between. The estimation of parameters is complicated by the nature of the model,

which only allows the coefficients corresponding to a given category to be estimated as a deviation from those of a reference category. To see this, we start by noting that the probability a child/household i will select occupational choice k is given by:

$$p_{ik} = \frac{\text{Exp}(Z_i \cdot \gamma_k + Y_{-i} \cdot \alpha_k + w_i \cdot \beta_k)}{\sum_j \text{Exp}(Z_i \cdot \gamma_j + Y_{-i} \cdot \alpha_j + w_i \cdot \beta_j)} \quad (7)$$

Taking ($j = 0$) as the reference state or occupation category, the equation above (7) can be re-written as:

$$p_{ik} = \frac{\text{Exp}[Z_i \cdot (\gamma_j - \gamma_0) + Y_{-i} \cdot (\alpha_j - \alpha_0) + w_i \cdot (\beta_j - \beta_0)]}{1 + \sum_{j=1}^2 \text{Exp}[Z_i \cdot (\gamma_j - \gamma_0) + Y_{-i} \cdot (\alpha_j - \alpha_0) + w_i \cdot (\beta_j - \beta_0)]} \quad (8)$$

$$\text{for } j = 1, 2 \text{ and } p_{i0} = 1 - p_{i1} - p_{i2}$$

Multinomial logit estimation permits the estimation of only the differences $(\alpha_j - \alpha_0)$, $(\beta_j - \beta_0)$, and $(\gamma_j - \gamma_0)$ for $j = 1, 2$. Since the transfer is conditional on the child being in state 1 or 2, the income variable $(Y_i + T_{ij})$ is asymmetric across alternatives. Thus in order to find the utility maximizing alternative (k^*), it is necessary to find estimates of the coefficients α_0 , α_1 and α_2 , for which the structural assumptions listed in equation 4 are useful.

Let $\hat{a}_j$ and $\hat{b}_j$ be the estimated coefficients of the multinomial logit model corresponding to the income and the child earning variables for alternatives ($j = 1$) and ($j = 2$), the alternative ($j = 0$) being taken as the reference. Using equation 4 we get the following system of equations:¹⁰

$$\hat{a}_1 = \alpha_1 - \alpha_0; \quad \hat{a}_2 = \alpha_2 - \alpha_0; \quad \hat{b}_1 = \alpha_1 \cdot M - \alpha_0; \quad \hat{b}_2 = \alpha_2 \cdot D - \alpha_0 \quad (9)$$

Which is equivalent to

¹⁰ To obtain equation 9, recall from equation 4 that: and that: for $j = 1, 2$.

$$\alpha_1 = \frac{\widehat{a}_1 - \widehat{b}_1}{1 - M}; \alpha_0 = \alpha_1 - \widehat{a}_1; \alpha_2 = \alpha_1 + \widehat{a}_2 - \widehat{a}_1; D = \frac{\widehat{b}_2 + \alpha_0}{\alpha_2} \quad (9a)$$

Using equation 9a, we can derive estimates of α_0 , α_1 , α_2 and D from the estimated coefficients of the multinomial logit and M ; where M is obtained from the estimation of equation 2. In estimating the residual terms $(v_{ij} - v_{i0})$, it is important to recall that these cannot be observed in a discrete choice model, and can instead only be known to belong to certain intervals. Thus the residual term for each child is drawn from the relevant interval, which is to say, in a way consistent with that child's observed choice of occupational category.¹¹

Finally, we note that equation 6 is not easily estimated without variable w_i , which is unobservable for children who are not in the labor market. The most rigorous approach would be to estimate the discrete choice model and the earning equation *simultaneously* by maximum likelihood techniques—a cumbersome procedure.¹² Correcting the estimation of the earning function for a selection bias turns out to be problematic as well (see discussion in BFL).¹³ Instead, the BFL model uses a simple approach, which has the advantage of transparency and robustness. This consists of estimating equation 2 by OLS, which is then used to predict the potential wage w_i for children who are not in

¹¹ For instance if child i has chosen state or category 1, the terms $(v_{ij} - v_{i0})$ must be drawn so as to satisfy the inequality.

¹² To handle simultaneously the random terms of the discrete choice model and that of the earning equation, a multinomial probit would then be preferable to a multinomial logit. This would then, however, pose the difficult challenge of integrating tri-variate normal distributions.

¹³ Proper identification using the inverse Mill's ratio to correct the earnings equation for selection bias requires variables/instruments which influence earnings but not the occupational/schooling choice of children. Such variables are not readily available from the data. Moreover, the standard correction using a two stage procedure is even weaker in the case of more than two choices.

the labor market. A random term u_i is added to the predicted earning of each non-working child to account for unobserved heterogeneity, by drawing from the distribution generated by the residuals of the OLS.

Simulating the impacts of a CCT program

Using the estimation steps described above, the outcomes of a given CCT program can be simulated with equations 5 and 6, but with the additional step of introducing a “means test” to identify eligible beneficiaries, which is a feature in CCT programs like *PROGRESA* and *BDH*. The means test is represented by assuming that the transfer T is provided only if household income is less than or equal to a pre-determined threshold Y_0 . Taking into account both the means-test and the conditionality of child attending school, child (household) i would choose the state or occupational category that yields the maximum utility among the following alternatives:

$$\begin{aligned}
 U_i(0) &= Z_i \cdot \gamma_0 + \alpha_0 Y_{-i} + \beta_0 w_i + v_{i0} \\
 U_i(1) &= Z_i \cdot \gamma_1 + \alpha_1 (Y_{-i} + T) + \beta_1 w_i + v_{i1} \quad \text{if } Y_{-i} + M w_i \leq Y_0 \\
 U_i(1) &= Z_i \cdot \gamma_1 + \alpha_1 Y_{-i} + \beta_1 w_i + v_{i1} \quad \text{if } Y_{-i} + M w_i > Y_0 \quad (10) \\
 U_i(2) &= Z_i \cdot \gamma_2 + \alpha_2 (Y_{-i} + T) + \beta_2 w_i + v_{i2} \quad \text{if } Y_{-i} \leq Y_0 \\
 U_i(2) &= Z_i \cdot \gamma_2 + \alpha_2 Y_{-i} + \beta_2 w_i + v_{i2} \quad \text{if } Y_{-i} > Y_0.
 \end{aligned}$$

As explained earlier, the estimation of the multinomial logit regression, combined with the relationships shown in equation 9a, allows us to simulate the utility maximizing decision for household i to choose among the alternatives specified in equation 10. It is easy to see that the introduction of a transfer can induce households to move from occupational category 0 (no schooling) to category 1 or 2, and also from 1 to 2 if it were the case that the household qualifies for the transfer only when the child went to school *and* stopped working.

The framework in equation 10 can be used to simulate the impacts of a variety of CCT programs that are conditional upon schooling enrollment, allowing for both the means test and transfer to be dependent on individual or household characteristics. That said, important caveats or limitations apply to this framework, which are closely related to the assumptions set out at the beginning of this section. Firstly, the model cannot account for the effects of any upper limit on transfers going to a single household, which is a direct result of the model ignoring the possibility that decisions affecting multiple children in a household may be made simultaneously. Secondly, household income excluding the earnings of the child is treated as exogenous. This does not take into account the possibility of the means test affecting adult labor supply. While this can be a serious issue when eligibility is defined by actual (or observed) income, it is less so when eligibility is defined by a score-based “proxy” means test that attempts to reflect permanent income.¹⁴

Programs and Data Description

PROGRESA was launched in Mexico at a time when economic growth was found to be insufficiently pro-poor and existing safety net programs were seen as ineffective (Coady 2004). *PROGRESA* was designed to alleviate poverty in both the short and long run, by making monetary transfers to poor families while requiring that they invest in the human capital of their children. Renamed *Prospera* in 2014 after being renamed *Oportunidades* in 2003, *PROGRESA* provides monetary grants to selected families (usually to the mothers) conditional on

¹⁴ Typically the proxy means test score is determined by factors such as ownership of durable goods and assets like land, housing conditions, education and occupational status of household members, and demographic characteristics like number of children and dependency ratio—none of which are directly influenced by the decision of an adult member to participate in the labor market.

children being enrolled in and regularly attending school, and on all family members attending scheduled visits to health care centers.¹⁵

PROGRESA was first implemented during the first half of 1998 in 320 randomly assigned rural villages out of 506 villages pre-qualified for immediate participation.¹⁶ The remaining 186 villages were assigned as control villages and received their first monetary transfers in December 2000. Implementation of the program in a phased manner and random selection of program villages for the first phase facilitated rigorous evaluations of the program.¹⁷

The baseline survey (*Encuesta Évaluation de los Hogares* or ENCEL) for *PROGRESA* was collected in 1997 in all 506 pre-qualified villages, covering 24,077 households in seven states. *PROGRESA* administrators ran a follow-up survey over the same set of households once every six months, generating a rich household panel dataset.

Table 1.
PROGRESA Socio-Economic Survey (ENCEL)

Treatment Localities (320)		Control Localities (186)		Total Localities: (506)		
Households		Households		Households		
Eligible for <i>PROGRESA</i>	Non-Eligible for <i>PROGRESA</i>	Eligible for <i>PROGRESA</i>	Non-Eligible for <i>PROGRESA</i>	Eligible for <i>PROGRESA</i>	Non-Eligible for <i>PROGRESA</i>	Total
11,623	3,233	7,173	2,048	18,796	5,281	24,077

Source: Skoufias (2005).

¹⁵ For a more detailed description of *PROGRESA* see Skoufias and Parker (2001) or Parker and Skoufias (2000).

¹⁶ These villages were randomly selected out of 6,396 villages (4,546 considered to be in the treatment group and 1,850 in the control group) using probability proportional to size. Villages were identified on the basis of a community score based on census information on characteristics such as educational levels, occupational composition and housing conditions.

¹⁷ Further details are available in Skoufias and Parker (2001).

Many *ex post* evaluations of *PROGRESA* were conducted on the basis of these surveys, including Skoufias and Parker (2001), Parker and Skoufias (2000) and Schultz (2004). Overall results show an increase in the enrollment rates of both boys and girls at the primary and secondary levels, a reduction in the incidence of illness among children of age 0–5 years, an increase in the annual growth rate of children of age 12–36 months and a reduction in all poverty indices. *PROGRESA* is now covering around 5 million households (18 percent of the country's population) with a budget equivalent to 0.4 percent of GDP, contributing around 20 percent of the income of participating families.

In 2003, Ecuador started *Bono de Desarrollo Humano* (BDH), which updated and re-targeted the transfers of the pre-existing program *Bono Solidario*. Under BDH, beneficiary households receive grants of \$15 per month under the conditions that children of ages 6–16 years are regularly enrolled in school with an attendance rate of at least 80 percent per month and children of ages 0–5 years make scheduled visits to health centers. The monthly \$15 transfers account for only 7 percent of household expenditure (Schady & Araújo 2006). BDH coverage reached one million households (5 million people), covering 40 percent of the Ecuadorian population with an estimated cost of 0.6 percent of GDP in 2005.

The baseline survey, collected between June and August 2003 for the evaluation of the BDH program, was drawn from four of 22 Ecuadorian provinces around the country. Within all four provinces, *paroquias* (parishes) were randomly assigned. For each selected *paroquia*, a random household sample was selected for the baseline and followed up in a survey in 2005 (after program implementation), generating a panel sample of 1,306 households with 2,875 children of age 6–17 years. Sixty two percent of households were eligible for the program, while the rest were not eligible to receive any transfer in the first two years. Schady and Araújo (2008) found BDH to have a significantly positive impact on school enrollment and negative impact

on child work, some of which may be attributable to many households believing school attendance to be a prerequisite for receiving benefits even though the conditions were never enforced by program administrators.

Comparison between *Ex Ante* and *Ex Post* Estimates of Program Impact¹⁸

The BFL model as described above can be used to simulate school enrolment (and labor participation of children) in the presence of the CCT program, which can then be compared with the outcomes for the counterfactual—namely, the observed scenario in the absence of CCT. When the simulation is conducted with the treatment group—a sub-sample of the full baseline survey sample—the impact obtained may be considered as an estimate of the “average intent to treat” (henceforth AIT) because it simulates the impact of the CCT program on the enrollment rate among all eligible children, regardless of whether they accept the transfer or not. Even if a child is eligible to receive the transfer, if the amount of transfer is insufficient to move him or her from category 0 to 1 or 2, the household would not accept the transfer and consequently, the transfer is not allocated to the household in the simulation.¹⁹

The AIT estimator is then given by:

$$AIT^* = E[P^*|treatment\ group, T = 1, X] - E[P|treatment\ group, T = 0, X] \quad (11)$$

Where P^* is a dummy variable taking the value 1 if children were

¹⁸ The main findings of the section were also found to hold for the Nicaraguan *Red de Protección Social* (RPS) program. The RPS analysis is omitted here to save space, but results are available upon request.

¹⁹ In maximizing its utility, an eligible household will decide to not accept the transfer if $U(0) > U(1)$ even when $Y_{-i} + M_{wi} \leq Y_{\theta}$, and $U(0) > U(2)$ even when $Y_{-i} \leq Y_{\theta}$.

enrolled in school after receiving transfer T , and P is a dummy variable taking the value 1 if children are currently enrolled in school. AIT* provides an estimate of the average impact of the availability of the program to eligible households in treatment communities.

Two aspects of the *ex ante* AIT estimator must be noted. First, it assumes good implementation of the program in treatment communities. While the simulation accounts for eligible households opting to not receive the transfer and send their children to school, it does not account for “exclusion errors” in targeting whereby households eligible and willing to receive the transfer (complying with its condition of sending children to school) are left out of the program. Neither does it take into account other types of implementation flaws, including “inclusion errors” of non-eligible households receiving the transfers, lack of compliance with the conditions of the transfer, or schools being unable to accommodate increases in enrollments. Second, the *ex ante* AIT estimator does not take into account any time trend effects on outcomes on which impacts are measured, which implies that it is best compared with *ex post* AIT estimates using a difference in difference (DID) method that removes the time trend effect from estimated impact.

The impact of *PROGRESA*

We compare the simulated *ex ante* impact of *PROGRESA*, applying the BFL model on the baseline survey (from 1997—see above), with *ex post* results from Skoufias and Parker (2001). All income variables from the baseline survey are converted into November 1999 pesos, and transfer amounts corresponding to the second half of 1999 (see Appendix, Table A-1 for the amounts) are applied in the simulations.

The selected sample comprises only children who were of age 8–17 years at the time of the baseline survey in 1997. Given our interest in *ex ante* estimates of impact, data from the follow-up survey on the same

children is not utilized for the simulation. Of the 33,609 children in the baseline sample the simulation is applied to, around 30 percent were not attending school, while only around 4 percent were combining work outside the home with school. Children not attending school were on the average older and with a higher education level than those who were in school and not working. Households with higher dropout rates were not necessarily poorer, indicating the importance of child income for the household. Child earnings increased with age and girls were more likely to drop out than boys, but boys were more likely to work and attend school simultaneously (see Appendix, Table A-2 for all descriptive statistics). Average household size was similar (around seven members) for the three groups of children and more educated parents were more likely to have children in school.

Dropout rates increase sharply around the age when children are expected to complete primary schooling, rising from 7 percent at age 11 to nearly 17 percent at age 12 and increasing exponentially thereafter to reach nearly 82 percent at age of 17. Because of the vast differences in enrollment across age groups, Skoufias and Parker suggest that the sample of children of age 8–17 years must be divided into two groups in assessing the potential impact of the program: children of primary school age (8–11 years) and those of secondary school age (12–17 years).

The experimental design of *PROGRESA* ensures in theory that all possible sources of bias were evenly distributed among participants and non-participants, so that differences between treatment and control groups can be strictly attributed to program effects. However, different papers have argued that the two groups were not fully comparable despite randomization. Skoufias and Parker find some evidence of systematic differences between the two groups and propose that a double-difference (DID) estimator, which takes into account any pre-existing differences, is preferable in evaluating the impacts of *PROGRESA*.²⁰ To be consistent with their methodology, the BFL model is estimated for all boys and girls aged 8–17 years in the baseline

sample, and the results are presented separately for three age groups that make for six age-gender categories.²¹

As explained earlier, the *ex ante* AIT estimator using BFL (as in equation 11 above) can be compared directly with the *ex post* results from Skoufias and Parker (Tables 5 and 6), who use the sample of all eligible households and measure the direct effect of the “intent to treat,” regardless of whether they in fact received a transfer or not. The *ex post* AIT estimates should be seen as a lower bound of the impact on households that actually received treatment, since the observed impact also includes the effect of flaws in program targeting, which would dilute the impact on eligible households. These estimates are derived from a regression-based approach that yields DID estimate of the program’s impact in each round (of the follow-up surveys), which is net of any pre-program differences between treatment and control households and time trends in the outcome indicator.

Skoufias and Parker (in Figures 4a and 4b of their paper) show that the mean school attendance rate of both boys and girls were nearly identical for control and treatment villages at the outset of the program. For treatment villages, the enrollment rate among boys was higher than that of girls in the baseline year, where the difference is entirely a result of the higher dropout rate among girls of age 12–17 years (Table 2). For the age group 8–11 years, enrollment was high at around 94 percent for both boys and girls.

Table 2 shows the AIT *ex post* estimates of *PROGRESA* impact on school enrollment by age and gender (Skoufias and Parker, henceforth SP) as well as the *ex ante* AIT estimates using the BFL simulation model. The *ex ante* AIT (conditional) estimates compare quite well

²⁰ Also, Behrman and Todd (1999) found the null hypothesis of mean equality with respect to household characteristics between treatment and control group to be rejected more frequently than expected.

²¹ Details of the estimation, including model specifications and regression results, are available on request.

Table 2.

*Impact of Transfers on Children's Occupational Choice (1997-1999):
Actual and Counterfactual School Enrollment Rates for PROGRESA
Target Population*

		Baseline (pre-program)	<i>Ex post</i> AIT1	<i>Ex ante</i> AIT2 Conditional	<i>Ex ante</i> AIT2 Unconditional
Boys	8-17 y.o	74.5%	-	4.0% ** (0.4%)	-0.1% (0.7%)
	8-11 y.o.	93.8%	1.8% ** (0.7%)	0.0% (0.4%)	-0.1% (0.5%)
	12-17 y.o	57.5%	5.8% ** (2.1%)	5.9% ** (0.8%)	0.0% (0.8%)
Girls	8-17 y.o	69.4%	-	4.3% ** (0.7%)	-0.1% (0.4%)
	8-11 y.o.	93.9%	-0.3% (1.0%)	-0.2% (0.4%)	-0.1% (0.4%)
	12-17 y.o	47.9%	9.5% ** (2.2%)	6.6% ** (0.8%)	-0.1% (0.8%)

Note. Source is baseline Survey 1997 and Rounds 1-4; authors' calculation. Standard errors and standard deviations in parentheses. For *ex post* AIT, robust standard errors are shown, which account for clustering of individuals within villages. For *ex ante* AIT, standard deviations computed by bootstrap method.

¹ Skoufias and Parker (2001) - Table 6 (results for November 1999 reported). Coefficients reported are the marginal effects of *PROGRESA* program on the probability of attending school.

² Results from simulation on baseline data using BFL model.

** Significant at 5% level; *significant at 10% level.

with the *ex post* estimates. Both show the impact of *PROGRESA* on enrollment to be much higher among children of age 12–17 years than children of age 8–11 years, and higher among girls than boys in the 12–17 age group. While the *ex ante* and *ex post* estimates are almost identical for boys of age 12–17 and girls of age 8–11, the *ex ante* estimates understate the impact for girls of age 12–17 and boys of age 8–11. Boys of age 8–11 years are the only group for which the two types of estimates disagree on whether there is any significant impact or not—the *ex post* estimate show a small positive impact while the *ex ante*

estimate is not different from zero. The confidence intervals of *ex ante* and *ex post* estimates for the same group overlay each other to a large extent. One advantage of *ex ante* methods is that they allow for the simulation of hypothetical scenarios. In a hypothetical case where transfers are unconditional, namely provided to all potential beneficiaries regardless of whether they are enrolled in school or not, the simulated effect on enrollment rate is found to be zero for all age-gender groups.

Table 3.

*Impact of Transfers on Children's Occupational Choice (1997-1999):
Actual and Counterfactual Proportions of Children Working for
PROGRESA Target Population*

		Baseline (pre-program)	<i>Ex post</i> AIT1	<i>Ex ante</i> AIT2 Conditional	<i>Ex ante</i> AIT2 Unconditional
Boys	8-17 y.o.	23.0%		-2.6% ** (0.3%)	0.0% (0.5%)
	8-11 y.o.	6.7%	-1.1% (0.8%)	0.3% (0.4%)	0.0% (0.4%)
	12-17 y.o.	37.3%	-4.7% ** (2.2%)	-3.7% ** (0.5%)	0.0% (0.7%)
Girls	8-17 y.o.	9.7%		-0.7% * (0.3%)	0.0% (0.3%)
	8-11 y.o.	4.2%	0.0% (0.0%)	0.4% (0.3%)	0.0% (0.4%)
	12-17 y.o.	14.6%	-2.3% * (1.3%)	-1.3% * (0.5%)	0.0% (0.5%)

Note. Source is baseline Survey 1997 and Rounds 1-4; authors' calculation. Standard errors and standard deviations in parentheses. For *ex post* AIT, robust standard errors are shown, which account for clustering of individuals within villages. For *ex ante* AIT, standard deviations computed by bootstrap method.

¹ Skoufias and Parker (2001), Table 5 (results for November 1999). Coefficients reported are the marginal effects of *PROGRESA* program on the probability of children working.

² Results from simulation on baseline data using BFL model.

** Significant at 5% level; *significant at 10% level.

Table 3 shows the results of a similar exercise comparing *ex ante* and *ex post* AIT estimates of the impact of transfers on the proportion of children who are working outside the home. The *ex ante* estimates are again quite close to the *ex post* SP estimates of impact. Both methods indicate that *PROGRESA* has the impact of reducing the rate of child labor among 12–17 year olds, but not among 8–11 year olds—quite consistent with what was seen in terms of impact of enrollment using both methods. Among the 12–17 year olds, the program’s impact on child labor is larger for boys than for girls. For both boys and girls of the 12–17 year age group, the *ex ante* estimates understate the size of the impact relative to the *ex post* estimates. Consistent with what was seen for enrollment, the *ex ante* simulations with unconditional transfers show no impact on child labor, suggesting that the conditionality of the transfer was a crucial factor for the impacts.

While it is encouraging to find that the BFL model has good predictive power when compared to results from *ex post* evaluations with experimental design, the similarity of results may be driven by the restricted and homogenous nature of the baseline sample (ENCEL) on which the simulations are run. Moreover, from a practical point of view, a specialized baseline survey of program areas may not be available in all cases to conduct *ex ante* simulations. CCT programs are often set up quickly to utilize a brief window of opportunity, which may leave little time for baseline data collection. In such cases, a nationally representative household survey, which has much more heterogeneity across households than any baseline survey, is often the only source of data available to conduct simulations on.

Therefore, for methodological as well as practical reasons, it is useful to test the validity of *ex ante* predictions from the BFL model using the Mexican National Household Survey (ENIGH). This is done by estimating the key parameters of the BFL model (α_j ’s, M and D) for all children of age 12–17 years with ENIGH 1996 (the round that was the closest to the launch of *PROGRESA*), and comparing the estimated parameters with those estimated with a similar model specification for

the same age group of children from the ENCEL survey.

The estimated parameters from ENIGH do not differ significantly from those estimated with the baseline survey (Table 4). For all key parameters, the 95 percent confident interval generated by ENIGH and ENCEL overlap each other and are thus statistically similar. However, ENIGH parameters generate smaller program effects compared to the results in Table 1 because all the parameters are a little smaller in size than ENCEL-generated parameter estimates. These results confirm the findings of Leite (2007) for Brazil, which show that estimations based on a more heterogeneous sample such as the one surveyed in the national representative PNAD leads to smaller parameter estimates. By estimating the same set of parameters for a sub-sample of homogenous households, Leite found a much higher impact on the enrollment rate and reduction of child work.

Table 4.

Estimating Key Parameters from BFL model: Using Baseline 1997 and ENIGH 1996 Surveys

Boys and Girls	Baseline - ENCEL		ENIGH
	8-17 y.o.	12-17 y.o.	12-17 y.o.
α_0	0.025** (0.008)	0.021** (0.008)	0.018** (0.009)
α_1	0.025** (0.009)	0.021** (0.008)	0.017** (0.009)
α_2	0.025** (0.009)	0.021** (0.008)	0.018** (0.009)
M	53%** (2.02%)	54%** (3.09%)	61%** (5.01%)
D	65%** (14.8%)	76%** (12.2%)	76%** (9.9%)

Note. Source is baseline survey 1997 and ENIGH 1996. Standard Deviations (in parentheses) computed by bootstrapping.

**Significant at 5% level; *significant at 10% level.

The simulation results presented in tables 2 and 3 and the estimated parameters presented above can be sensitive to different model specifications, as is the case with any simulation model. Two different sets of model structure are used for exploring how variations in model parameters would affect key parameters of the BFL model, which are α_0 , α_1 , α_2 , M and D . First, 11 different models for equation 2 (standard Becker-Mincer human capital model) were used, while equation 4 (utility function model) was held constant. Subsequently, for each of the 11 models for equation 2, two versions of equation 4 were used. The average values of key parameters for both sets of model structures are not significantly different from those estimated by the original model, which indicate that the main results of the simulations would hold for a wide range of specifications of the model.²²

The BFL model also makes the estimation of program effects on poverty and inequality possible, which is done by setting 750 pesos as the maximum transfer per household (consistent with the rules of *PROGRESA*), using the same poverty line as in Skoufias and Di Maro (2008),²³ and taking into account individual level decisions through the BFL model. The proportion of poor households in the *baseline treatment* village sample is estimated to be 61 percent, which is the same as in Skoufias and Di Maro. Applying the DID estimator on poverty measures, Skoufias and Di Maro find that between November 1997 and November 1999, *PROGRESA* was responsible for a 16.5, 24 and 27 percent reduction in the poverty headcount rate, poverty gap and squared poverty gap, respectively, in treatment villages. Our *ex ante* simulations suggest a similar impact of *PROGRESA* on poverty headcount rate (13 percent reduction), but a somewhat larger impact for

²² Appendix, Table A-3 shows the average values of parameters and their standard deviations for both sets of model structures in comparison with the parameters from the selected model presented earlier in Table 4.

²³ The poverty line is equivalent to 176.17 pesos per month (for details, see Skoufias & Di Maro 2008)

poverty gap (27 percent reduction) and squared poverty gap (33 percent reduction), as shown in Table 5.²⁴

Table 5.
Poverty Index: Observed (Ex Post) and Simulated (Ex Ante)

	FGT(0)		FGT(1)		FGT(2)	
	%	σ	%	σ	%	σ
Treatment communities (baseline)	60.9%	1.1%	30.8%	1.1%	21.5%	1.1%
Change in poverty ex post (Skoufias & Di Maro 2008) ¹	-16.5%	1.6%	-24.3%	1.5%	-29.2%	1.4%
Change in poverty ex ante (simulated) ²	-13.1%		-26.6%		-33.1%	

Note. Source is baseline Survey 1997 and Rounds 1-4. Prices as of November 1999 (pesos). Poverty line: mean of Nov 98 consumption per capita (taken from Skoufias & Di Maro 2008)

¹ Skoufias and Di Maro (2008), Table 4;

² Authors' calculation using BFL model.

The final step in our analysis is the simulation of the cost of the program based on the BFL model. Table 6 shows that 4,567 of 7,110 (or 64 percent of) poor families in treatment villages would enroll in the program and accept the condition of school attendance. The average cash transfer for these families in treatment villages is estimated at 301 pesos per month, which is very close to the average transfer of 300 pesos in January 2002 for 1.6 million families observed from official program data. By scaling up the cost of the program from the BFL model to the national figure using ENCEL 1997, the total cost of the program is estimated to be around 6.08 billion pesos, which is very close to the actual program cost of 6.06 billion pesos in January 2002 from the official program data (Table 6).

²⁴ The *ex ante* results in Table 5 remain the same when the simulations are done without incorporating the conditionality of schooling. This suggests that conditionality does not matter for meeting the poverty reduction objectives of a CCT, which is consistent with intuition.

Thus the BFL model applied to the *PROGRESA* program in Mexico appears to have strong predictive power for the impact of the program on outcome indicators (enrollment rate, child labor and poverty) as well as the cost of the program, when compared against relevant benchmarks—namely AIT impacts estimated by *ex post* evaluations using experimental methods and actual or observed program cost. Since *ex post* AIT estimates represent a lower bound of the impact on households that actually *received* treatment (for reasons explained earlier), the *ex ante* estimates of program impact from BFL should also be seen as a lower bound for the actual impacts of the program on beneficiary households.

Table 6.
Estimated and Observed PROGRESA Cost

	BFL Simulation ^{1,2}	January 2002 ^{1,3}
Number of families with children receiving benefits	4,567	1,681,254
Average monthly transfer for families with children	\$31	\$300
Estimated total annual transfer	\$1,17,580	\$6,061,221,243
Scaling up total transfer: simulated average times families in 2002	\$6,080,632,364	n.a.

¹ November 1999 prices.

² Estimated by BFL model using baseline survey of 1997, rounds 1-4.

³ National official numbers from: http://www.oportunidades.gob.mx/indicadores_gestion/ene_feb_02/indice.htm

The Impact of *Bono de Desarrollo Humano* (BDH)

As before, we compare the *ex ante* estimates of program impact from the BFL model with *ex post* findings from Schady and Araújo (2008) (henceforth SA). For *ex ante* simulations, we use the baseline survey of BDH collected in Carchi, Imbabura, Cotopaxi and Tungurahua (described above), which is also used by SA as the baseline data in analyzing impact. By the authors' own admission, the

results from SA should be treated with some caution as they are not based entirely on random assignment.

A summary of descriptive statistics from SA reveals no significant differences between the pre-defined treated and non-treated samples for a large number of variables (see Appendix, Table A-4), suggesting that the random design of the experiment in terms of the intended treatment and control groups was successful. However, a comparison between recipients and non-recipients of BDH transfers shows significant differences between the two groups for variables like child enrollment, parental education and household size (Appendix, Table A-4). Thus selection into the BDH program appears to be non-random.

In order to avoid misspecification due to bias from the contamination of both treatment and control groups, SA computed the “treatment-on-the-treated” or TT estimator using the Two Stage Least Square method. The TT estimator suggests that the BDH increased enrollment rates among compliers by 9.7 percent, with the statistic being only weakly significant (i.e. significant at the 10 percent level but not at 5 percent). Notably, around 25 percent of households in the sample believed that compliance with the condition of school attendance was mandatory for receiving the cash transfer (to be called “conditioned households” henceforth), despite the fact that this conditionality was not enforced by BDH administrators. Out of these conditioned households, 55 percent were from treatment and 45 percent from control groups. Following the re-weighting scheme of Hirano et al. (2003), SA reported the Difference-in-Difference or DD estimator, adding a dummy variable for “conditioned” and an interaction term of the “conditioned” dummy variable with the treatment group dummy variable.

As previously mentioned, our BFL model estimator is a type of “intend-to-treat” (ITT) estimator. In this case, it is easy to transform the *ex post* TT estimator from SA (results from Table 3 in SA) for comparison with our estimator, as follows:

$$TT = \frac{ITT}{\Pr\{\text{receive BDH} \mid \text{Treated}\}} \text{ or } ITT^* = TT \cdot \Pr\{\text{receive BDH} \mid \text{Treated}\} \quad (12)$$

Another issue to take into account is the amount of transfer to be allocated to each child. Our BFL model is based on the assumption that the decision to enroll (or not) every child in school is made (by the child or the household) independently, taking into account the transfer amount for the child. But the BDH program transfers \$15 per household, independent of the number of enrolled children in the household, which makes the decision of sending or keeping every child in school essentially a household level decision.

If we set \$15 as the direct transfer amount, each child will add to its utility an amount of \$15. α_j (for all $j = 1, 2$),²⁵ which will lead to an overestimation of the effects of the program. This can then be considered as the upper bound of the estimated impact of BDH from the BFL model. The BFL model does, however, allow us to construct alternative scenarios of how the \$15 is distributed within the household, namely the size of the de facto transfer to every child when there are multiple children in the household. We simulate three alternative scenarios. The first is one where the transfer received by each child is equal to \$15 divided by the number of children aged 6–17 years in the household, with a minimum of \$3 per child. In the second scenario, we set an individual transfer of \$5 because the average number of potential beneficiaries is close to three per household. In the third scenario, the individual transfer is set at \$7.50. In order to account for the cost of the program and average size of transfer per household, we set the household maximum to the official amount of \$15 and also compute the impact *without* any household maximum, for every scenario.

Table 7 shows all *ex post* estimators obtained by SA and the *ex ante* results for the different scenarios using BFL. As expected, the *ex ante* or simulated ITT estimator under the assumption of \$15 transfer

²⁵ Refer to equation 4 earlier.

per child (ITT¹), at 14.6 percent, is much larger than the 10 percent *ex post* impact measured by SA. Setting a household maximum of \$15, the annual cost of the program under scenario 1 (estimated at \$166,000) is a useful benchmark. The effect on enrolment generated by scenario 1 is implausible, as a child living with its siblings will never make a decision on program participation based on a \$15 transfer.

The next three scenarios have lower estimated impacts ranging from 5.9 percent to 9 percent; and for all three cases, the *ex post* ITT estimator falls in the 95 percent confidence interval of the *ex ante* ITT estimator. For these scenarios, the average transfer per household ranges from \$9.80 to \$15.95. The estimated cost of the program, regardless of whether the household maximum is set at \$15 or not, never exceeds the benchmark cost under scenario 1 with \$15 per household. These results suggest that larger program effects would have been obtained from making individual transfers of \$7.50 per child so that the amount of transfer per household ranges from \$7.50 to \$45 depending on the size of the household.

For the reasons described earlier, we cannot calculate *ex post* results for the BDH program in a way that would allow for a direct comparison with *ex ante* predictions as we have done for the case of *PROGRESA*. However, the predictive power of the BFL model is consistent with our expectations conditional on the assumption made about the allocation of the household transfer among beneficiaries. As we would expect, considering each child in the sample to be a household under the BFL definition (and receiving \$15) overestimates the impact of the BDH. Keeping the average transfer per household close to or above \$15 and using smaller amounts of transfer per child, the BFL model estimates program effects on enrollment rate that are not statistically different (at 5 percent level of significance) from *ex post* results reported by SA for all simulated ITT estimators.

Conclusion

The results of our paper suggest that *ex ante* and *ex post* methods for estimating impacts can effectively complement each other to inform the design and evaluation of social programs. In principle, combining an *ex ante* simulation exercise using a structural model of schooling choices with an experimental evaluation design would help policymakers to assess the implications of alternative choices for program design, potential impacts and cost implications upfront, and rely on *ex post* evaluation results to validate and identify the actual impacts of the program. For *ex ante* models to be used this way, however, their results need to be credible in terms of predictive power. Our findings make some progress in this direction—the results suggest that *ex ante* predictions of certain impacts of CCT programs using the BFL model match up well with *ex post* “gold standard” results from experiments.

The primary analysis of the paper consists of comparing the *ex ante* simulated impact on enrolment rates using the BFL model with the *ex post* impact estimates from a randomized experiment. For the case of the *PROGRESA* program in Mexico, we can also compare the *ex ante* and *ex post* estimates of program impact on poverty. Such a comparison of methods is also a test to validate the hypothesis central to such *ex ante* models: cross-sectional income effects, estimated on the basis of the representative household sample, coincide with the income effects generated by a program.

For *PROGRESA*, we estimate that the predicted impact tends to understate, but are not significantly different (at 5 percent level of significance) from, *ex post* “intent to treat” or ITT estimators of impact for all age groups of boys and girls. Moreover, the BFL model estimates impacts on poverty indices that are comparable with those from an *ex post* evaluation. For the Ecuadorian program *Bono de Desarrollo Humano* (BDH), the use of the BFL model, with reasonable assumptions about how the household level transfers would translate to

de facto transfers for every child, yields simulated impacts that are again statistically similar to *ex post* ITT impact estimates. The fact that *ex ante* predictions from the BFL model are validated by *ex post* estimates also lends confidence to the use of the model for analyzing other types of questions related to program design, such as the likely impact of *PROGRESA* on enrollments if it were an unconditional cash transfer. Using the BFL model we find that transfers without conditions would *not* have produced the enrollment impacts seen from the conditional transfer. In the case of BDH, the *ex ante* simulations suggest that program impact on enrollment rates could have been higher if the transfers were set at an amount per child (instead of a flat amount for every household), for an unchanged overall cost of the program.

Thus the *ex ante* method proposed by Bourguignon, Ferreira and Leite (2003) can be a powerful simulation method in designing or reforming CCT programs that aim to improve enrollments. One major advantage is that in many cases, such a model can allow for comparison between the likely impacts of alternative ways of designing a program—without the time, cost and logistics involved in setting up multiple pilot experiments to obtain *ex post* evaluation results. There are important caveats as well to the use of the BFL model, the most important among which is that *ex ante* analysis cannot incorporate any unanticipated changes that may occur during the process of program implementation, be that in the implementation process or in any other extraneous events that can affect the behavior of program participants. The *ex ante* exercise would then need to be followed up with an *ex post* impact evaluation that is able to capture the *actual* effect of the program on the outcomes of interest, taking into account all the possible changes in the program and the environment in which it was being implemented.

Table 7.

Children's Occupational Choices in Ecuador 2003-2005: Actual and Counterfactual Enrollment Rate for Target Population

	Ex post				Ex ante			
	d1	TT (a)	Probability of receiving benefits if eligible (b)	ITT (c)=(a)(b)	Simulated ITT ¹	Simulated ITT ²	Simulated ITT ³	Simulated ITT ⁴
Table 4 – Results 2SLS	10.00% 4.72%	10.00% 4.72%	77%	7.7%	14.58% 0.53%	7.63% 0.82%	5.86% 1.10%	9.00% 0.60%
Average transfer per household with ceiling of \$15 per household					15.00	11.91	9.80	12.31
Number of households					645	631	596	603
Annual cost					116,100.00	90,192.00	70,080.00	89,100.00
Average transfer per household without household ceiling					32.33	11.92	10.35	15.95
Number of households					645	631	599	605
Annual cost					250,200.00	90,264.00	74,400.00	115,830.00

Note. Source is Schady and Araújo (2008), table 4 column 3. Standard Deviation computed by bootstrap method.

¹ R\$15 per child assuming unitary child household.

² \$15 divided by number of children age 6-17 in the household. Minimum transfer \$3 per child.

³ \$5 per child.

⁴ \$7.50 per child.

** Significant at 5% level; * significant at 10% level.

References

- Alderman, H. & Yemtsov, R. (2014). How Can Safety Nets Contribute to Economic Growth? *World Bank Economic Review*, 28(1), 1-20.
- Attanasio, O., Meghir, C. & Santiago, A. (2005). *Education choices in Mexico: Using a structural model and a randomized experiment to evaluate PROGRESA* (IFS Working Paper EWP0501). London: Institute for Fiscal Studies (IFS).
- Behrman, J. & Todd, P. (1999). *A report on the sample sizes used for the evaluation of the education, health and nutrition program (PROGRESA) of Mexico*. Washington, DC: International Food

Policy Research Institute.

- Bourguignon, F. & Ferreira, F. (2003). *Ex Ante* Evaluation of Policy Reforms using Behavioral Models. In F. Bourguignon & L. da Silva, (Eds.), *Economic policies on poverty and income distribution: Evaluation techniques and tools* (Chapter 6). World Bank: Washington DC; New York, NY: Oxford University Press.
- Bourguignon, F., Ferreira, F., & Leite, P. (2003). Conditional cash transfers, schooling, and child labor: Micro-simulating Brazil's *Bolsa Escola* program. *World Bank Economic Review*, 17(2), 229-254.
- Cardoso, E. & Souza, A. (2004). *The impact of cash transfers on child labor and school attendance in Brazil* (Working Paper No. 04-W07). Nashville, TN: Department of Economics, Vanderbilt University.
- Coady, D. (2004). *Alleviating structural poverty in developing countries: The approach of PROGRESA in Mexico* (Background Paper for the 2004 World Development Report). Washington, DC: World Bank. Retrieved from: http://econ.worldbank.org/files/27999_Coady.pdf
- De Janvry, A. & Sadoulet, E. (2006). Making conditional cash transfers more efficient: Designing for maximum effect of the conditionality. *World Bank Economic Review*, 20(1), 1-29.
- Fiszbein, A., Schady, N., Ferreira, F., Grosh, M., Keleher, N., Olinto, P., & Skoufias, E. (2009). *Conditional cash transfers: Reducing present and future poverty*. Washington, DC: World Bank.
- Garcia, M. & Moore, C. (2012). *The cash dividend: The rise of cash transfer programs in Sub-Saharan Africa*. Washington, DC: World Bank.
- Grosh, M., Del Nino, C., Tesliuc, E., & Ouerghi, A. (2008). *For protection and promotion: The design and implementation of effective safety nets*. Washington, DC: World Bank.
- Heckman, J. & Hotz, J. (1989). Alternative methods for evaluating the impact of training programs. *Journal of the American Statistical*

Association, 84(408), 862-880.

- Heckman, J., Ichimura, H. & Todd, P. (1997). Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme. *Review of Economic Studies*, 64(4), 605-654.
- Heckman, J., Ichimura, H. & Todd, P. (1998). Matching as an econometric evaluation estimator. *Review of Economic Studies*, 65(2), 261-294.
- Hirano, K., Imbens, G., & Ridder, G. (2003). Efficient estimation of average treatment effects using the estimated propensity score. *Econometrica*, 71(4), 1161-1189.
- LaLonde, R. (1986). Evaluating the econometric evaluations of training programs with experimental data. *The American Economic Review*, 76(4), 604-620.
- Leite, P. (2007). *Les politiques d'allocations scolaires et l'inégalité des chances au Brésil*. Mimeo.
- Lise, J., Seitz, S. & Smith, J. (2005). *Evaluating search and matching models using experimental data* (IZA Discussion Paper No. 1717). Bonn: Institute for the Study of Labor (IZA). Retrieved from: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=789464
- McKenzie, D., Gibson, J. & Stillman, S. (2006, April). *How Important Is Selection? Experimental vs. Non-Experimental Measures of the Income Gains from Migration* (IZA DP No. 2087). Bonn: Institute for the Study of Labor (IZA). Retrieved from: <http://ftp.iza.org/dp2087.pdf>.
- Parker, S. & Skoufias, E. (2000). *The Impact of PROGRESA on Work, Leisure, and Time Allocation* (Report submitted to PROGRESA). Washington, DC: International Food Policy Research Institute.
- Ravallion, M. (2008). Evaluating anti-poverty programs. In T. Schultz & J. Strauss (Eds.), *Handbook of development economics*. Amsterdam: North-Holland.
- Schady, N. & Araujo, M. (2006). *Cash transfers, conditions, school enrollment, and child work: Evidence from a randomized experiment in Ecuador* (Policy Research Working Paper No.

- 3930, Impact Evaluation Series No. 3). Washington, DC: World Bank.
- Schady, N. & Araujo, M. (2008). Cash transfers, conditions, and school enrollment in Ecuador. *Economía*, 8(2): 43-70.
- Schultz, T. (2000). School subsidies for the poor: Evaluating the Mexican *PROGRESA* poverty program. *Journal of Development Economics*, 74(1), 199-250.
- Skoufias, E. (2005). *PROGRESA and its impacts on the welfare of rural households in Mexico* (Research Report 139). Washington, DC: International Food Policy Research Institute.
- Skoufias, E. & Di Maro, V. (2008). Conditional cash transfers, adult work incentives, and poverty. *Journal of Development Studies*, 44(7), 935 - 960.
- Skoufias, E. & Parker, S. (2001). Conditional cash transfers and their impact on child work and schooling: Evidence from the *PROGRESA* program in Mexico. *Economía*, 2(1), 45-86.
- Smith, J. & Todd, P. (2005). Does matching overcome LaLonde's critique of nonexperimental estimators? *Journal of Econometrics*, 125, 305–353.
- Todd, P. & Wolpin, K. (2006). Assessing the impact of a school subsidy program in Mexico: Using a social experiment to validate a dynamic behavioral model of child schooling and fertility. *American Economic Review*, 96(5), 1384–1417.

Appendix

Table A-1.
PROGRESA *Monthly Cash Transfer*

Educational grant per child		
	Boys	Girls
Primary		
3rd	80	80
4th	95	95
5th	125	125
6th	165	165
Secondary		
1st	240	250
	250	280
2nd		
3rd	265	305
Cash grant for food per household		125
Maximum transfer per household		750

Source: Skoufias and Parker (2001). Figures in pesos - July to December 1999.

Table A-2.

Descriptive Statistics from PROGRESA Baseline Survey, Children aged 8-17 in 1997

	Not Attending School	Working and Attending School	Attending School	Total
Age	14.8	12.6	11.2	12.3
Years of Schooling	5.5	5.3	4.6	4.9
Household per capita income	301.7	238.5	241.3	259.3
Observed children's earnings				
8	536.5	282.1	0.0	381.0
9	484.2	217.3	0.0	301.6
10	593.2	317.1	0.0	424.8
11	680.4	389.4	0.0	537.9
12	706.1	323.1	0.0	518.8
13	618.3	332.9	0.0	541.9
14	702.8	415.1	0.0	653.4
15	793.8	453.2	0.0	760.8
16	813.5	549.5	0.0	792.6
17	838.7	690.5	0.0	830.4
Years of schooling of more educated parent	2.7	3.3	3.7	3.4
Age of older parent (years)	47.5	45.3	45.2	45.9
Number of household members	7.3	7.3	7.2	7.2
Male (%)	48.2	72.8	51.5	51.4
Father works in agriculture sector (%)	46.6	34.7	51.3	49.2
Proportion of universe (%)	30.0	4.3	65.7	100
Population	10,081	1,435	22,093	33,609
8 (%)	4.9	2.5	92.6	100
9 (%)	4.2	3.5	92.3	100
10 (%)	5.6	3.5	90.9	100
11 (%)	7.0	4.5	88.6	100
12 (%)	16.6	6.1	77.3	100
13 (%)	28.1	5.8	66.1	100
14 (%)	42.0	5.4	52.6	100
15 (%)	57.7	4.1	38.2	100
16 (%)	71.3	4.3	24.5	100
17 (%)	81.6	3.1	15.2	100

Source: Baseline Survey 1997; and Authors' calculation.

Table A-3.

Estimating Key Parameters from BFL Model for Different Model Structures using Baseline 1997 Surveys and Different Model Specifications

Boys and Girls (8-17 y.o.)	Selected Model 1 ¹	Variations in Equation 2 Model Specification 2 ²	Variations in both Equation 2 and Equation 4 Model Specifications 3 ³
0	0.025 ** (0.00845)	0.036 ** (0.00810)	0.033 ** (0.00805)
1	0.025 ** (0.00845)	0.035 ** (0.00770)	0.032 ** (0.00806)
2	0.025 ** (0.00846)	0.036 ** (0.00773)	0.033 ** (0.00806)
M	0.530 ** (0.02016)	0.534 ** (0.00394)	0.534 ** (0.00394)
D	0.653 ** (0.14763)	0.712 ** (0.05325)	0.755 ** (0.13666)

Source: Baseline Survey 1997; Authors' calculation.

** Significant at 5% level; * significant at 10% level.

¹ Selected Models and Standard Deviation computed by bootstrap method.

² Average parameters and parameters Standard Deviation for 11 different model specifications on Equation 2-standard Becker-Mincer human capital model.

³ Average parameters and parameters Standard Deviation for 22 different model specifications for both Equation 2-standard Becker-Mincer human capital model, and Equation 4-utility function model.

Table A-4.

Differences between Treatment and Control Groups – BDH Baseline Survey

	Differences between Treatment and Control Groups		Differences between recipients and non-recipients of the BDH transfers	
	Mean: Control	Difference	Mean: non-recipients	Difference
Probability that child is enrolled	0.770 (0.421)	-0.003 (0.017)	0.736 (0.441)	0.052*** (0.018)
Mean years of schooling completed	4.59 (2.47)	0.031 (0.086)	4.65 (2.46)	-0.071 (0.090)
Probability that a child is working	0.463 (0.499)	0.012 (0.027)	0.478 (0.500)	-0.014 (0.027)
Child age	11.67 (2.86)	0.047 (0.094)	11.81 (2.86)	-0.174 (0.097)
Child is male	0.517 (0.500)	-0.031 (0.018)	0.506 (0.500)	-0.009 (0.019)
Log of per capita expenditures	3.44 (0.543)	0.013 (0.030)	3.48 (0.564)	-0.046 (0.031)
Asset index	0.007 (0.882)	-0.015 (0.049)	0.016 (0.966)	-0.028 (0.052)
Father's education	4.72 (2.58)	0.087 (0.156)	4.44 (2.76)	0.523*** (0.160)
Mother's education	3.52 (2.99)	0.283 (0.168)	3.30 (2.99)	0.598*** (0.172)
Household size	5.83 (1.78)	-0.116 (0.102)	5.64 (1.86)	0.208** (0.105)
F-statistic (p-value)		0.15		0.002

Note: Source is Schady and Araújo (2006) – Table 2. All means refer to baseline values. Standard errors in estimated difference for child-specific variables adjust for within-sibling correlation. **Significant difference at the 5 percent level; *** at the 1 percent level. The sample is limited to households who had school-aged children in both the baseline and follow-up surveys. Sample size is 2876 for all child-specific variables, and 1309 for all household-specific variables except father's education (n=1227) and mother's education (n=1279).