International Journal of Humanities
and Social Sciences (IJHSS)
ISSN(P): 2319-393X; ISSN(E): 2319-3948
Vol. 6, Issue 3, Apr-May 2017; 91-100

International Academy of Science,
Engineering and Technology
Connecting Researchers; Nurturing Innovations

EX ANTE IMPACT ASSESSMENT OF CONDITIONAL CASH TRANSFER PROGRAMS USING AN AGENT-BASED MODEL

SONIA ESPERANZA MONROY[1] & HERNANDO DIAZ[2]

[1]Professor, Department of Industrial and Systems Engineering,
Faculty of Engineering, Universidad Nacional de Colombia, Bogotá, Colombia

[2]Research Scholar, Department of Electrical and Electronic Engineering,
Universidad Nacional de Colombia, Bogotá, Colombia

ABSTRACT

A model for *ex ante* impact evaluation of Conditional Cash Transfer Programs (CCT) is described. Assessment of the impacts of such programs is critical for the design of public policies. We developed an agent-based model, aimed at predicting the decisions that families would make, given their characteristics and environment but without any assumption of rationality. Thus, expected behaviors and effects can be predicted, based on the individuals' traits, by computing the odds of making different choices concerning children´s education. The model we developed has the capacity to simulate the decision making process, including adaptive behaviors and learning strategies to realistically represent actual situations in more detail.

The model was applied to Brazil´s Bolça Familia program in 2011 and results compared to a previous study. Our model indicates that this Conditional Cash Transfer program has been successful and that, as a consequence, school attendance has increased during the period between the evaluations.

KEYWORDS: Agent-Based Model, Conditional Cash Transfer, Ex Ante Impact Assessment, Public Education Policy

INTRODUCTION

Currently, there is a consensus on the importance of evaluating the impact of public policies, given the size of the investments required and the need for their replicability. Considering the limited investment resources, it is essential to select the policies and programs that produce the desired impacts on the population.

Conditional cash transfers (CCTs) have become an important public policy instrument in developing countries. They are, as (Fizbein and Schady 2009) describe them, "programs that transfer cash, generally to poor households, on the condition that those households make prespecified investments in the human capital of their children". Most existing CCT programs require that the prespecified investments be made in the children´s education, health or nutrition. These programs have been implemented with varying success in several Latin American countries, including, Brazil, Mexico, Colombia and Ecuador. Many countries have invested substantial amounts of money in order to reduce inequalities, by enhancing the odds of poor children succeeding in life. Therefore, it is important to assess how effective those investments have been, since otherwise the resources allocated to these programs could be used in other, perhaps more effective ways.

Impact assessment differs from the mere assessment of results of projects or programs. Often, the analysis of results refers to short-term effects of a decision of investment or financing. Impact assessments, on the other hand, analyze the medium and long-term effects, both positive and negative, caused by the intervention resulting from policies, programs or projects, OECD (2002).

Generally, impact assessments are done *ex post*, i.e., after the program has been carried out. However, in recent years it has become clear the need to make impact assessments of social policy before its implementation (*ex ante* evaluations) in order to select the policies that produce the greatest possible positive effect under a given budgetary constraint. Alternatively, decision makers might be trying to design a policy that guarantees a given effect, while minimizing its cost. Actually, it is desirable to be able to predict different effects of the policy on the agents of the system, to choose the program to be implemented, instead of undertaking social programs whose actual overall effects are not known accurately.

Ex ante impact assessments aim to predict the effect on society of financing programs or projects and, in general, of investment policies. The predictions are based on the simulation of the effects of program funding, through general economic equilibrium models or through deterministic and stochastic simulation models. For this purpose, models aiming to predict the macroeconomic outcome of economic decisions have been developed. They are based on models of the behavior of individuals who make decisions while seeking to maximize their welfare. These models give an estimate of the sum of all individual decisions and the resulting effect on the economy. Based on those models, *ex ante* impact assessments can be obtained. Behavioral models allow decision makers to forecast the impact of policies. Then, an optimization process may be used to select the most effective policy in each situation.

More recently, a different kind of models for ex ante evaluations have been considered that allow much more detail in the representation of the individuals making decisions. These methods consider a large number of individual economic agents whose characteristics resemble those of the population being studied. Both individual-level variations and fluctuations due to varying geographical or social conditions can be incorporated into the simulation models. Currently, these assessment models, based on simulation, are of two kinds: micro simulations and agent-based models. We use an agent-based model (ABM) in this work.

Microsimulation has been defined as “a technique used to model complex real life events by simulating the actions and/or impact of policy change on the individual units (micro units) that make up the system where the events occur” Harding (2007). This type of model has been used extensively in economics. Applications include population dynamics to evaluate effects of employment policies or pension reform schemes. There are also applications in the general urban and economic planning.

Leite, P., A. Narayan, and E. Skoufias, (2011), carried out a comparison between the impacts, estimated through microsimulation models of programs of conditional cash transfers with the actually measured values (*ex post*) in randomized experiments in Mexico (Progresa program) and Ecuador. The results show that the predicted values by an *ex ante* micro-simulation approach agreed within reasonable accuracy with the *ex post* data obtained by experimental and quasi-experimental methods. In a similar way, Bourguignon et al (2003) evaluated the impact of a program of conditional cash transfers in Brazil: the Bolça Escola program. In this project cash stipends were awarded to families whose children, between 5 and 15 years of age, attended school regularly.

The authors developed a model of utility function used to simulate decision-making within a family, based on information on family income, family characteristics and the child. We have used the Bolça Escola program's sequel to illustrate the application of the model.

An alternative approach to impact assessment, originated independently in several areas of knowledge (ecology, artificial intelligence, computer networks, etc.), is one in which the behavior of individuals is simulated independently. This is the so-called Agent Based Model approach (ABM). These are computation-intensive simulation models in which individual behavior of a given population is studied by simulating the behavior of independent entities (agents) represented in the simulation by rules or decision functions. This allows greater flexibility since agents can assume any kind of behavior, not necessarily determined by economic conditions. In addition, agents can communicate through mechanisms provided in the environment where the dynamics evolves. And it is also possible to model the existence of subsystems with coordination functions.

In agent-based models as well as in micro simulation, the population under study is modeled by a synthetic population whose characteristics are similar to those of the actual population. The object is to study the behavior of individuals comprising the population under analysis. Large sets of agents are considered, chosen in a way that they are representative of the population at large. Often, the entire population is represented, thus leading to models with millions of individuals. This means that it is necessary to have large amounts of information, depending on the study. Models of the decision-making process are also included, as in econometric models. However, in micro simulations and ABMs, the behavioral models are used individually. Thus, there is a model for each individual that makes up the population.

Recently, agent-based models have been used to assess the impact of social and agricultural programs. In particular, they have been utilized to model intervention policies in agricultural markets. Nevertheless, agent-based models have been used much less frequently than other alternatives. Filipski and Taylor (2012), for instance, used an agent-based model to study the impact of income subsidy policies for rural populations in Ghana and Malawi. The impact of direct payments to residents was studied, in comparison to alternative mechanisms of subsidizing inputs, production and other welfare policies.

METHODOLOGY

Economic models usually rely on general assumptions about the *rationality* of decisions made by economic agents. And, while this premise may be justified under perfect or near perfect market conditions, there are many situations where actual decisions are made in a way that is far from rational. Therefore, the model to be described next relies only on empirical distributions of the actual decisions made by the people in a sample. Although a utility function may be implicit, its purpose is purely theoretical and it only serves to identify the variables that are relevant to the household´s welfare. These are the variables that may affect the choices made by a family regarding their children´s school attendance. We present in this section the basis for the development of the model. The assumptions made are presented first and then the procedure for the simulation is described.

In Conditional Cash Transfer (CCT) programs, stipends are assigned to families. Therefore, it is considered that it is a joint family decision whether to comply with the conditions (e.g., sending the child to school instead of letting him get a job) or not. Hence, it is necessary to estimate the probability of making each of several possible choices regarding the child's human capital.

Assumptions of the Model

We assume that the families make decisions related to the CCT based on social, regional and economic conditions as well as on the amount of cash received as stipend. In order to develop a model that may be used for ex ante evaluations, it is important to include the conditions of the CCT and its amount as part of the set of variables that predict the odds of making each choice. A decision variable is introduced that describes feasible choices regarding the child's participation in the CCT program. It is also assumed that the expected value of the wage that a child might earn under his conditions may be weighted against the prospective cash transfer from the program. Therefore, an estimate of the wage y_i earned by a child (i) with some characteristics is considered to be available. In practice, this value must be obtained from a sample.

It is also assumed that information on family and child characteristics is available. The information is summarized in a vector Z_i that contains variables that characterize the children, like age, school grade, birth order. Also included are household characteristics like the location, family size, age and educational level of elder family member, etc. The monetary income is considered separately in another variable Y_i that includes all family earnings except the money received as stipend or the child's wages.

Next, a multinomial logistic model is utilized to establish the relation between the variables identified as potentially affecting the choices and the probability of making each of the possible choices. A multinomial logistic regression (Agresti and Kateri, 2011) is thus used to establish these relationships. Let's assume that the possible choices regarding the participation of the household in the program are given by a discrete variable X_i describing the decision made by the i-th child's family. Possible choices are $X_i \in \{0, 1, \ldots, n\}$. Then, a multinomial logistic model is fitted that gives the odds of choosing option j versus option k, using standard methods.

$$\log\left(\frac{p_j}{p_k}\right) = Z\,\delta_j + (Y + y)\beta_j \quad j \neq k$$

.

Then, the probability of choosing each option is computed for each individual family.

Finally, an agent-based model is used to simulate the decision that each family makes, given their social, geographic and familiar conditions. Taking into account that no utility function is used explicitly, an agent-based model implementation was preferred. This opens the way to more complex representations of the family decisions, including adaptive behaviour. It is also possible to study the influence of migrations and other geographically relevant dynamics.

Procedure

The approach we have used to estimate the impact of the cash transfer programs can be summarized in the following stages:

- Find a source of information on the variables to be included in the model. Our model requires that information, on the actual choices made by the individual decision-makers with respect to their participation in the program, be available. In the example to be presented, this means the decision on sending a child to school, or to work, or both is known for a sample of the population.
- Generate a synthetic population representing as closely as possible the statistical characteristics of the real population being studied: number, gender composition, rural and urban distribution, education level, etc.

- Estimate the available monetary resources for the family, resulting from each possible choice made. In the case we study, this means that the expected income of every child who wants to work is estimated as a function of the child´s characteristics and other local variables like the median income for the state where his family resides and whether the family lives in a rural or urban neighborhood, among others.
- Estimate the odds of making each different choice by using a multinomial logistic regression estimation of the likelihood of family decisions, based on a sample. When the actual choices are known for the individuals included in the survey being used, the multinomial logistic regression allows us to compute the probability of each possible choice for each individual.
- Simulate the effect of the policy for different scenarios, using the agent-based model, in order to compare the potential impacts of the various policy alternatives. A number of simulation *replications*[1] will be necessary to ensure reliable computation of mean values and confidence intervals of impact variables.

CASE STUDY: BRAZIL'S "BOLÇA FAMILIA PROGRAM – BFP"

The **Bolça Familia Program** (BFP) in Brazil provides a monthly stipend in cash to each family with children from ten to seventeen years of age that satisfy certain conditions. A sum of money will be given for each child in the age range, provided that they attend school on a regular basis and regular medical check ups. The objective is to reduce poverty and inequality in the country by encouraging school attendance, better nutrition and health.BFP emerged in 2003 from the integration of several previous social programs. It provides a monthly subsidy between R$35 and R$105 to qualifying families having a monthly household income up to R$137. Total annual cost of the program was around 0.4% of Brazil's gross domestic product at that time.

A study database provided by the Brazilian Institute of Geography and Statistics (IBGE) in its annual survey entitled "National Survey by Household. Sample of households" (PNAD) of 2011 was taken as a base for our analysis. Survey income data were obtained for persons over 10 years of age in the 27 states that make up the country. One of these data is the income received from the person´s main job in Reals (R$). However, only 7.7% of this population has income information in the sample, leaving out a large fraction of children.

A predecessor of this program had been studied in 2003 by Bourguignon and Ferreira. While testing our model, we wanted to compare our results to their analysis.

Estimation of Prospective Children' S Wages

Household income varies widely from one state to another. Therefore, a revenue base that takes these disparities into account is necessary to estimate children's wages. The state median income was chosen as representative of the income level for the state. ABecker - Mincer human capital type model was developed to predict how the wages of children behave in their entirety taking as baseline the wages of children who actually earned it. With the data provided by the PNAD of 2011 and based on the work of Bourguignon and Ferreira (2003) we used the following equation:

[1]Since a stochastic simulation is being used, each run will be a different realization of the underlying stochastic process. Thus a number of replications should be averaged to account for random fluctuations.

$$\log y_i = \beta Z_i + \alpha X_i + \mu_i \quad (1)$$

Where the logarithm of wages of the *i*-th child observed, y_i, is expressed in terms of a vector $Z_i$of characteristics of the child i, such as age, education, sex, etc. and of a decision variable X_i indicating the family decision with respect to the child's school attendance.

$$X_i = \begin{cases} 0 & \text{Child works; he does not study} \\ 1 & \text{Child works and studies} \\ 2 & \text{Child attends school, exclusively} \end{cases}$$

To predict the wages, a linear regression using ordinary least squares was performed to estimate the parameters: the vector β and the coefficient α, using the statistical package Stata. Only the data for children between 10 and 17 years of age, working during the week referenced in the survey PNAD of 2011 who had a declared income from their main job, were considered. The model shows that the observed income grows with age. That behavior agrees with economic theory: the older and more educated the person, the higher the expected wages are, for the economically active population. The distribution of predicted revenues also presents a growing trend, and although its magnitude is slightly greater than those observed, the characteristic we focus our attention on is the trend. The coefficients of our model are similar to those obtained by (Bourguignon 2003), and the *p*-values show that the variables are statistically significant.

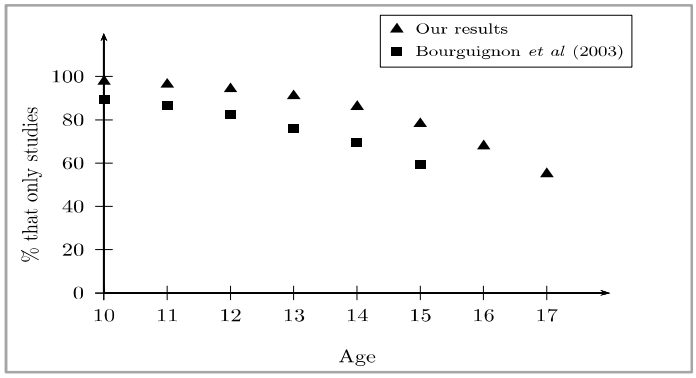
Estimation of the Odds of Family Decisions

It is possible to estimate the probability f the person who makes the decisions in the family choosing one of the three options available for the X_i variable, based on a sample of families who actually made those decisions. In order to estimate the odds that a household with given economic and family conditions make a given choice, a multinomial logistic regression was done using 51,815 observations of children between 10 and 17 years old. The results are not presented here, due to space limitations, but they show that the estimated income variable described above has a positive marginal effect when the decision was made to work exclusively, and an opposite effect when it was decided that the child should attend school and work. This result makes sense if the child that only works is more likely to generate higher income compared to one who divides his time between study and work.

RESULTS

The first and most important result is that 97% of children between 10 and 15 years are attending school. This means that the program could be considered successful since one of its primary objectives is to ensure the children´s school attendance. This result is even more promising when compared with previous results from Bourguignon and Ferreira (2003), in which for the period of 1999 the total number of children attending school was 93.9%. The number of children attending school decreases to 92.77% when considering the total group of children up to the age of 17 years.

Simulations show a marked decreasing trend of the percentage of students who attend school with increasing age. While for children younger than 13, the percentage of children who only attend school is higher than 90%, this number is significantly reduced for ages from 15 years to reach 55% at the age of 17, as shown in Figure 1.To make a comparison, they are presented in the same axes as the percentages calculated for 2009 by Borguignon and Ferreira (2003). In this figure, the triangles represent the estimated values using our model for the period 2011, while the squares are the results of the micro simulations made by Bourguignon. Our results include the ages of 16 and 17, which were not included in the original BolçaEscola program but were incorporated into the new BFP.



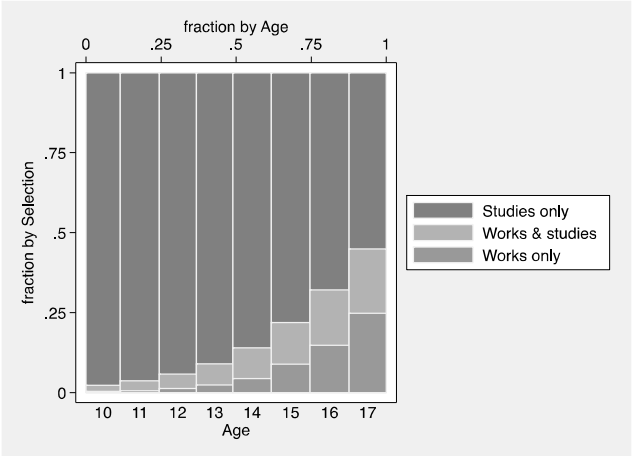
Source: (Bourguignon *Et Al* 2003) and Authors' Calculations

Figure 1: Fraction of Children Attending School and Not Working, By Age

Even though it is not possible at this time to establish whether the conditions under which our model was developed are consistent with the model of Bourguignon, a simple comparison of the percentages of young people who have made the decision to attend school instead of working shows that those numbers have grown significantly for all ages. In fact, the current percentage of 15 year olds who attend school without working is higher than the percentage of children of 13 years of age who were in the same situation during the previous period.

These results suggest that, at least from the point of view of discouraging school dropouts, the Bolça Escola and Bolça Familia programs have had a significant impact and should be considered a success. Nonetheless, it is still worrisome that almost 25% of 17 year olds (24.83%) is dedicated to work, not to study.

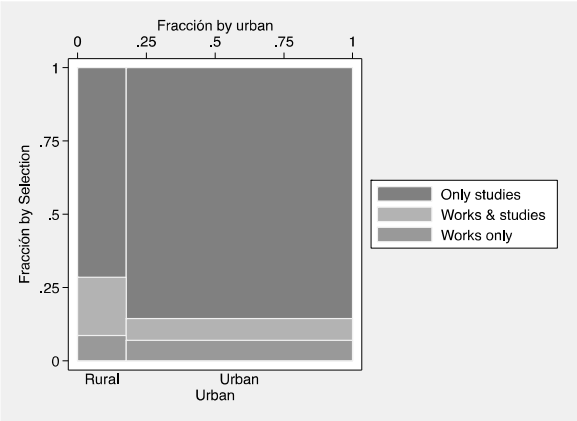
Figure 2 summarizes the overall results, by age, for the study population. There is a clear association between older age and the probability of dropping out of school. To try to quantify this trend, a Kruskal-Wallis γ statistic (seeAgrestiand Kateri (2011)) was computed. This statistic is an attempt to give numerical values to this qualitative relationship. In the case of association between age and relative importance of education with respect to revenue (quantified by X_i), the estimated value is $\gamma = -0.6053$.This shows a strong (negative) association between the two variables, although the relationship is clearly not linear, as shown in Figure 2. That graph shows the strong positive association between age and the likelihood of dropping out of school.



Source: Authors' Calculations based on PNAD 2011

Figure 2: Fraction of Children of a Given Age in Each Condition

Additionally, we are interested in looking at the influence of other variables on the decisions a family makes about the education of their children. First, we wanted to study whether there were differences between the decisions that the families living in urban areas and those living in rural areas made about their children's education. It was found that there are indeed differences and a value of $\gamma = 0.3652$ was found. The relationship between these variables is best seen in Figure 3.



Source: Authors' Calculations based on PNAD 2011

Figure 3: School Attendance Decisions by Rural or Urban Living Area

Although the association is not as strong as with age, there is a moderately strong association of school attendance-related decisions with the area where the family resides. It is noticeable, for example, that for the rural population, the percentage of youth who only study is 71.5% while in urban areas that value is 85.64%. It would be interesting to try to assess whether the causes are structural, such as greater distance to school or greater relative contribution to the family finances or whether it is a matter of perception. It is clear that the fundamental difference is among those who study and work, simultaneously. In this case, the fraction for rural areas is 2.7 times that for urban areas. An analysis by age shows that the difference grows as age increases.

The analysis also shows that more educated parents do imply less likelihood of the child starting to work early. The difference is not as pronounced among those who have completed only primary school or who completed only a few years of high school. However among those having completed high school the difference in attitude with respect to early age work is quite different from those less educated. Results were also studied from a gender perspective: the fraction of children who do not attend school but only work is approximately the same for both sexes. And this applies to all age groups. However, the fraction of children who study and work at the same time is significantly higher for boys than for girls in the same situation. A more detailed age analysis shows that the difference is practically constant (the fraction of boys who work and study is nearly twice that for girls), for all age groups.

CONCLUSIONS

We used an agent-based model to assess the impact of investments on social programs. It has been shown that it is possible to evaluate the impacts of a given program using a model calibrated with actual data from the program. A model of this type allows the evaluation of the program in an *ex ante* form by predicting the outcomes and impacts of a given investment program. Thus, decision makers can assess several alternative investment programs before choosing one. Therefore, an optimal design of public policies becomes possible.

The model has been applied to *ex ante* evaluation of the impact of Brazil's Bolça Familia program. Publicly available data from a PNAD survey in 2011 were used to calibrate the model. This database, although of great breadth and helpful in the preparation of this work, was short on point estimates of children's revenue, an information required to generate the marginal effects of a multinomial logit model.

The results of this evaluation show that Brazil's Bolça Familia program has been successful from the point of view of encouraging children´s school attendance. At least this is what can be concluded from a comparison of the results of this work with the results of Bourguignon and Ferreira (2003).

Agent-based models allow us to overcome some of the weaknesses of traditional econometric models, like the assumption of rational decision-making. More realistic behaviors, e.g., adaptation and cheating, may be simulated with agent-based models.

These models can be profitably applied to the assessment of impacts of social investment programs as long as the required information is available. At the present moment, a general model of this type is being incorporated into a tool for *ex ante* evaluation of the impact of research program funding.

REFERENCES

1. Agresti, A. and Kateri, M. (2011). *Categorical data analysis* (pp. 206-208). Springer Berlin Heidelberg.
2. Bourguignon, F. and Ferreira, F. (2003). *Ex ante evaluation of policy reforms using behavioral models. The impact of economic policies on poverty and income distribution*, 123-142.
3. Filipski, M. and Taylor, J. E. (2012). A simulation impact evaluation of rural income transfers in Malawi and Ghana. *Journal of Development Effectiveness*, 4(1), 109-129.
4. Fiszbein, A.and Schady, N. R. (2009). *Conditional cash transfers: reducing present and future poverty*. World Bank Publications.
5. Harding, A. (2007). Challenges and opportunities of dynamic microsimulation modelling. In *Plenary paper presented to the 1st General Conference of the International Microsimulation Association, Vienna* (Vol. 21).
6. Leite, P. G., Narayan, A. and Skoufias, E. (2011). How do ex ante simulations compare with ex post evaluations? evidence from the impact of conditional cash transfer programs. Evidence from the Impact of Conditional Cash Transfer Programs (June 1, 2011).*World Bank Policy Research Working Paper Series*, Vol.
7. OECD Development Assistance Committee. (2002). *Glossary of key terms in evaluation and results-based management. Organisation for Economic Co-operation and Development*, Paris.