Randomization and Social Program Evaluation: The Case of Progresa
Susan W. Parker and Graciela M. Teruel

The online version of this article can be found at:
http://ann.sagepub.com/content/599/1/199

Published by:
SAGE
http://www.sagepublications.com

On behalf of:
American Academy of Political and Social Science

Additional services and information for The ANNALS of the American Academy of Political and Social Science can be found at:

Email Alerts: http://ann.sagepub.com/cgi/alerts

Subscriptions: http://ann.sagepub.com/subscriptions

Reprints: http://www.sagepub.com/journalsReprints.nav

Permissions: http://www.sagepub.com/journalsPermissions.nav

Citations: http://ann.sagepub.com/content/599/1/199.refs.html

>> Version of Record - Mar 31, 2005

What is This?
In this article, the authors analyze the development of Progresa, a Mexican antipoverty program that conditions monetary transfers to human capital investment of its beneficiary families. The program is the principal antipoverty strategy of the Mexican government and has served as a model for similar programs in a number of countries in Latin America. Progresa is also noteworthy because it was subject to a rigorous evaluation effort in rural areas, which included an experimental design. The authors first describe the rationale behind the design of Progresa, in particular, the conditioning of monetary transfers to children’s school attendance and regular health clinic visits. The authors then turn to the evaluation effort, analyzing the randomized trial, the evaluation results, and the effect of the evaluation on the evolution of the program. Finally, the authors consider the limitations of the evaluation as well as areas where more research is needed.

Keywords: randomization; experimental design; Progresa; Mexico

Introduction

We analyze in this article the development of Progresa (the Education, Health and Nutrition Program), a Mexican antipoverty program that conditions monetary transfers to human capital investment of its beneficiary families. The program is the principal antipoverty strategy of the Mexican government, representing in 2003, 46.5 percent of the federal government’s annual poverty budget. The program began operating in 1997 in poor rural areas and has recently expanded to cover urban areas. By the end of 2003, more than 4.2 million families were receiving benefits, with the goal of 5 million families by the end of 2004. The program has captured attention at the international level and

NOTE: The authors gratefully acknowledge support for the development of this article from the Campbell Collaboration. They also benefited from comments received at the Campbell Collaboration/Rockefeller Foundation Conference on Place-Randomized Trials in Bellagio, Italy (2002) and in New York City (2003).

DOI: 10.1177/0002716205274515
has served as a model for similar programs in several other countries in Latin America. The Inter-American Development Bank recently gave Mexico its largest loan ever with the objective of extending Progresa to urban areas. Leading economists have also praised its success and advocated its expansion in other contexts (see Becker 1999; and Krueger 2002).

The program, while providing cash transfers, is aimed at increasing families’ investment in human capital as defined by education, health, and nutrition. To achieve this objective, Progresa conditions cash transfers on children’s enrollment and regular school attendance as well as clinic attendance. These transfers correspond on average to a 22 percent increase in the income levels of the beneficiary families and are given directly to the mother of the family. The program also includes in-kind health benefits and nutritional supplements for children up to age five and pregnant and lactating women.

The program is noteworthy because it was subject to a rigorous evaluation effort in rural areas that included an experimental design. A subset of communities eligible to receive Progresa was randomly assigned to a treatment (320 communities) or control (186 communities) group. For the approximately twenty-four thousand households living in these communities, household interviews were carried out both prior and after program implementation over a period of three years. An organization, the International Food Policy Research Institute, was hired to provide an external evaluation on the impacts of Progresa. These impacts included direct impacts (e.g., education, health, and nutrition) as well as other potential impact indicators, including child and adult work, consumption patterns, women’s status, and transfers.

In what follows, we describe and analyze the design of Progresa, in particular, the rationale for conditioning monetary transfers to children’s educational attendance and regular health clinic visits and the fact that benefits are given directly to the woman head of the household. We then turn our attention to the evaluation effort, analyzing the randomized trial, the evaluation results, and the effect of the evaluation on the evolution of the program. Finally, we consider the limitations of the evaluation and areas where more research should be focused, for instance,

---

Susan W. Parker (Ph.D.), professor of economics at Centro de Investigación y Docencia Económicas, A.C. (CIDE) in Mexico City, is an expert in program design and family behavior. Since its inception, she has played a central role in the evaluation of Progresa, a large-scale nationwide cash transfer program that now serves as a model for many countries around the world. She has extensive fieldwork experience including innovative work tracking movers within Mexico and to the United States. Her recent work has been published or is forthcoming in the International Economic Review, the Journal of Population Economics, and the Review of Economic Development.

Graciela M. Teruel (Ph.D.) is professor of economics at Universidad Iberoamericana in Mexico City. She is an expert on poverty and household welfare in Mexico and is a leading expert in the design of innovative research surveys. She codirects the Mexican Family Life Survey, an extremely innovative, ongoing, longitudinal survey of individuals, households, families, and communities in Mexico. Her recent work has been published in the Journal of Development Economics, Economics and Human Biology, and the Journal of Population Economics.
whether the grant amounts represent the “correct” amounts (in terms of maximizing impact for a given cost), the effect that sample attrition may have on the evaluation results, and the implications of the relatively short duration of the experimental design.

A Description of Progresa

Progresa began operating in 1997, as a program aimed at poor households living in marginalized rural communities. The program was at first implemented in communities with fewer than 2,500 inhabitants, although a small number of communities with between 2,500 and 14,999 inhabitants (called semiurban) were also included. Urban communities, defined as areas with 15,000 inhabitants or more, began to be incorporated in the year 2001.

Table 1 shows the incorporation of beneficiary families in Progresa by geographic area over time through the end of 2003. Currently, the majority of beneficiary families are found in rural areas. Semiurban areas account for almost seven hundred thousand families, and the remainder, about five hundred thousand, are found in urban areas.
It is important to note that the same general package of benefits is provided in both rural and urban areas. We now turn to a description of these benefits.

**Specific benefits**

Perhaps the first innovation of Progresa is the fact that it combines three different components, that is, education, health, and nutrition, in one program. The reason for linking these three components in one program, as stated in the original document proposing Progresa (Poder Ejecutivo Federal 1997), was based on the philosophy that the interactions between the components would enhance the effectiveness of an integrated program over and above the separate benefits from each component. For instance, children who suffer from malnutrition are more likely to drop out of school or repeat years of school, implying that attempts to ensure that children go to school will be more effective if combined with adequate nutrition and health programs, thereby helping to break the vicious cycle of intergenerational poverty transmission (Poder Ejecutivo Federal 1997).

Under the education component, Progresa currently provides monetary educational grants for each child younger than twenty-two years of age enrolled in school between the third grade of primary and the third grade of high school. It is important to note, however, that grants were originally only provided through secondary school. The addition of education grants at the high school level was implemented in the year 2001. These grants (Table 2) increase as children progress to higher grades and, beginning at the secondary level, are slightly higher for girls than for boys. The amounts of the monthly grants range from about US$10.50 (105 pesos)
in the third grade of primary to about US$58 (580 pesos) for boys and US$66 (660 pesos) for girls in the third year of high school.

The second component, health, provides basic health care for all members of the family, with a particular emphasis on preventive health care (Table 3). These services are provided by public health institutions in Mexico. The third component, nutrition, includes a fixed monetary transfer equal to about $US15.50 (155 pesos) monthly for improved food consumption, as well as nutritional supplements, which are principally targeted to children between the ages of four months and two years and pregnant and lactating women. They are also given to children aged two to four if any signs of malnutrition are detected.

Perhaps the first innovation of Progresa is the fact that it combines three different components, that is, education, health, and nutrition, in one program.

Receipt of the benefits is contingent on fulfillment of certain obligations by the beneficiary families. The monetary educational grants are linked to the school attendance of children so that if a child misses more than three school days in a month (for unjustified reasons), the family will not receive the grant that month. Similarly, families must complete a schedule of visits to the health care facilities in order to receive the monetary supports for improved nutrition (Table 4).
school attendance of children and family health visits are verified through school and clinic records. Progresa has a maximum limit of monthly benefits for each family currently equivalent to about $US90 for families with children in primary and secondary school and $US160 for those with children in high school.

All of the monetary benefits are given directly to a woman of the family, most frequently the mother. Benefits are provided directly to the female beneficiary by wire transfer in offices and modules that are installed nearby the communities. In some urban areas, Progresa is using bank accounts to transfer money to its beneficiaries.

The targeting mechanism differs between rural and urban areas. In rural areas, Progresa selects its beneficiary families through a three-stage targeting mechanism. In the first stage, Progresa uses geographic targeting to select poor regions and communities. Communities with a high level of marginality, as measured by aggregate socioeconomic community characteristics, are selected to participate in the program. Next, to identify the beneficiary families, Progresa carries out a survey of socioeconomic conditions for all households in the selected communities. With this data, discriminant analysis is used to identify beneficiary households from nonbeneficiary households. Households are evaluated to be in extreme poverty not just on the basis of income but on the basis of a number of other characteristics as well, such as dwelling characteristics in the household, dependency ratios, ownership of durable goods, animals and land, and the presence of disabled individuals.

Once beneficiary households have been identified, an assembly is arranged in the community where the list of selected families is made public and an agreement is reached among all families in the community. It is worth noting that in practice, this last step rarely resulted in significant changes to the list of beneficiary families.

### Table 4

**Annual Frequency of Health Care Visits Required by Progresa**

<table>
<thead>
<tr>
<th>Age Group</th>
<th>Frequency of Checkups</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Children</strong></td>
<td></td>
</tr>
<tr>
<td>Younger than 4 months</td>
<td>3 checkups: 7 and 28 days, and at 2 months</td>
</tr>
<tr>
<td>4 to 24 months</td>
<td>8 checkups: 4, 6, 9, 12, 15, 18, 21, and 24 months, with 1 additional monthly weight and height checkup</td>
</tr>
<tr>
<td>2 to 4 years old</td>
<td>3 checkups a year: 1 every 4 months</td>
</tr>
<tr>
<td>5 to 16 years old</td>
<td>2 checkups a year: 1 every 6 months</td>
</tr>
<tr>
<td><strong>Women</strong></td>
<td></td>
</tr>
<tr>
<td>Pregnant</td>
<td>5 check-ups: prenatal period</td>
</tr>
<tr>
<td>During purporeum and lactation</td>
<td>2 checkups: in immediate purporeum and 1 during lactation</td>
</tr>
<tr>
<td><strong>Adults and youths</strong></td>
<td></td>
</tr>
<tr>
<td>17 to 60 years old</td>
<td>1 checkup per year</td>
</tr>
<tr>
<td>Older than 60 years old</td>
<td>1 checkup per year</td>
</tr>
</tbody>
</table>
In urban areas, the targeting mechanism is substantially different, and there is an element of self-selection in that potential beneficiaries must show up to a Progresa module in their geographic area of residence (local advertising is used to make potential beneficiaries aware of the opportunity to receive Progresa). At the module, their basic socioeconomic levels are assessed; for those that pass this initial qualifying test, a home visit is programmed to verify socioeconomic information; and based upon this information, a similar discriminant analysis as in rural areas is used to decide whether the household is eligible for Progresa.

An Analysis of the Design of Progresa

Progresa represents a significant change in the provision of social programs in Mexico. In this section, we analyze in detail some of the more innovative design features of the program, in particular those relating to (1) the conditionality of the benefits to human capital investment and (2) gender aspects.

Human capital and conditionality

Among the most important aspects of the program is its focus on investment in human capital. Whereas Progresa was designed with the objective of providing some immediate monetary assistance to families in extreme poverty, an equally important objective is that of increasing the human capital of its beneficiaries in the future. In this way, Progresa sees the lack of human capital in the poor population as a central cause of poverty and hopes that by increasing investment in human capital, above all, in children, the future poverty of these individuals (e.g., when they become adults) will be reduced. In accordance with the views of most economists, Progresa considers that the returns to investing in human capital, as measured by education, health, and nutrition are important. Greater investment in human capital is expected to have permanent effects in increasing future income of beneficiaries (Becker 1993; Mincer 1974). The theory of human capital, and the measurement of “returns to education” or the lifetime increase in income that can be expected through an increase in schooling years has been verified in numerous empirical studies in developing and developed nations. Furthermore, recent studies have indicated that both health and nutrition also have important returns in terms of future income (Strauss and Thomas 1995, 1998). Thus, economic theory and empirical economic studies were clearly instrumental in the design of Progresa.

One might ask why poor families do not invest “enough” in human capital. Progresa’s answer is that families may be aware of the benefits of human capital investment but cannot afford to invest as much as they would like (Poder Ejecutivo Federal 1997). The costs of school attendance are not just monetary costs associated with attending school but the opportunity costs of sending children to school. The opportunity cost of sending children to school is the income or value of income that children would earn if they were working rather than attending school. Since
families need this income for current consumption, they take their children out of school at early ages and send them to work. Progresa substitutes for the income children would earn if they were working and thus, theoretically at least, permits parents to send their children to school. Again, this is a vision clearly influenced by economists (see Becker 1993).

Progresa requires children to miss no more than three days a month or lose the education grant in that month. This conditionality is one aspect that has received most attention at the international level. However, the rationale for conditioning benefits to human capital investment is not clearly stated in the original document explaining the motivation for Progresa. If the problem is that parents do not have the current income to finance their children's school enrollment, this would imply, in the words of economists, a credit constraint. And while it might imply that the policy of giving monetary transfers is appropriate, it is not obvious why the transfers should be conditional on behavior such as sending children to school. Many economists might argue that conditionality requirements are paternalistic (e.g., the implication is that government knows better than the parents how they should invest and allocate resources to their children) and therefore not necessarily welfare improving relative to unconditional transfers. Under this scenario, it might be preferable (e.g., better for the family) to give unconditional transfers. One potential motivation for conditioning benefits is the assumption that there are some social returns to investing in education that are not reaped by the individual, for example, social returns are higher to private returns. If individuals only take into account private returns when deciding their (or their children's) investment in human capital, they will underinvest relative to the social optimum. A review of the literature, however, was unable to find empirical studies that demonstrated that social returns were significantly higher than private returns. Martinelli and Parker (2003) explored alternative reasons why a welfare-minded government might con-

[The decision to send monetary benefits to the mother] was motivated by the social science literature that argues that resources under the control of women tend to have a greater impact on the well-being of children than resources under the control of men.
dition transfers, focusing on conditionality as a way to implement outcomes closer to the mother's preference in the context of intrahousehold bargaining.

A final point related to the conditionality is that the program assumes that the returns to education are likely to be high, even in the relatively isolated areas where Progresa began operating and where it is likely that school quality is low. Resources such as libraries or computers are notably lacking in schools in rural areas. Many communities do not even have secondary schools within easy commuting distance (Coady and Parker 2004). Many of the secondary schools in rural areas are *telesecundarias*, which provide televised classes, as well as a single instructor who helps children with exercises in all different materials. In contrast, urban areas have general secondary schools, with a different teacher for each class material. It is dubious whether the returns to schooling are as high in rural areas as in urban areas. This is an area that has not been sufficiently studied. A study of school quality and the returns to education as well as other forms of human capital would be extremely valuable in terms of predicting the long-term effects of the program in rural Mexico.

**Gender aspects**

In addition to the emphasis placed on prenatal and postnatal care, Progresa has two gender aspects that have received international attention. The first is the point that the monetary benefits are received by the mother of the household. This design feature, as stated in program documents, was motivated by the social science literature that argues that resources under the control of women tend to have a greater impact on the well-being of children than resources under the control of men. Again, this is a literature dominated predominantly by economists (see, e.g., Thomas 1990). At the time Progresa was developed, there were no studies in the context of Mexico that would validate the relevance of this assumption. The implicit assumption was that Mexico is similar to Brazil and other countries for which empirical evidence did exist. Recent studies (Rubalcava, Teruel, and Thomas 2004; and Attanasio and Lechene 2002) support the point, in Mexico as elsewhere, that resources under the control of women have greater positive impacts on children than resources under the control of men.

The second important gender aspect is that, beginning with the education grants at the level of junior high, the grant levels are higher (about 10 to 15 percent) for girls than boys. This feature was originally motivated, according to Progresa, by the observation that in rural areas, girls tended to have a higher dropout rate than boys after finishing primary school. Thus, the higher grant levels were aimed at compensating for this lower achievement.

Nevertheless, as shown in Behrman, Sengupta, and Todd (2000a) and Parker and Federzini (2001), actual attainment of girls in terms of years of completed schooling in rural areas is quite similar to that of boys. The seeming paradox between lower enrollment and similar attainment in years of schooling can be explained by the point that boys tend to have higher repetition rates than girls. That
is, when girls enroll in school, they tend to do better than boys, that is, finish school levels faster.\textsuperscript{2}

The Evaluation of Progresa

The randomized trial

We turn now to a description of the randomized trial that Progresa carried out as part of its evaluation strategy. It is worth emphasizing that the evaluation of Progresa was quite novel in Mexico in many respects. Progresa was the first social program in Mexico to carry out a rigorous evaluation of program impacts. Program officials emphasized from the beginning of the program the importance of developing a rigorous evaluation; in general, it represented a strategy to try to ensure that the program would not be eliminated with changes in the government. As part of the strategy to provide credible results of program impacts, Progresa officials hired the International Food Policy Research Institute (IFPRI) to carry out an extensive external evaluation of program impacts between 1998 and 2000 (Skoufias 2001).

One of the advantages of the Progresa evaluation is that it was planned in conjunction with the program. The evaluation was implemented at an early stage, a point where the program was not yet so large that it would be difficult to construct a control group for a randomized evaluation.

The evaluation design was heavily influenced by the works of economists in the field of evaluation, who have generally argued that a randomized design, although not without potential difficulties, can have great advantages over nonexperimental methods in obtaining unbiased estimates of program impact (Heckman 1992). The Progresa evaluation can be considered to be somewhere in between a randomized experiment and a quasi-experimental evaluation. The general principle behind the evaluation was to use randomization to the extent that it was feasible. In this case, for a sample of communities eligible to participate in Progresa, the evaluation strategy was to carry out a randomization between control and treatment communities.

Initially, in the case of Progresa, randomization at the household level was considered but judged to be infeasible by program officials. The reasoning here was that randomizing households within communities, in a context where communities are quite small and “everyone knows everyone” would have created problems between those in the treatment group and those in the control group. Program officials feared that it would create the perception of discretionality in the selection process (e.g., households equally poor in the community where some are receiving benefits and others are not) (Gomez de Leon, Parker, and Hernandez 1999).

The alternative chosen was to select a sample of Progresa communities and randomize communities between the treatment group and the control group. This was a perhaps natural level at which to randomize as it fits well with the actual selection.
mechanism of Progresa beneficiaries where in the first stage localities are selected before passing to a second stage where households within communities are determined to be eligible. A potential limitation of randomizing at the community level, however, was the likelihood of observing less randomness at the individual level (discussed below). Behrman and Todd (1999) noted that randomizing at the community level implied that one was unable to analyze locality-specific treatment effects. This is the case if one uses a dummy variable to represent the impact of Progresa; nevertheless, if one uses amounts received by Progresa as the impact variable, one can obtain locality-specific impacts.

Progresa selected a sample of 506 communities in seven states as its “evaluation” sample, which corresponded to 24,077 households. These seven states were among the first states to receive Progresa benefits. Of the 506 communities, 320 communities correspond to the treatment group and 186 correspond to the control group. The 320 treatment communities were randomly selected using probabilities that were proportional to the size distribution of communities from a universe of 4,546 communities in the seven states. Using the same method, a sample of 186 control communities were drawn from a sample of 1,850 communities within the seven states that were scheduled by Progresa to receive benefits at a later time. Treatment beneficiary household began to receive benefits in May 1998, whereas control households began to receive benefits in December 1999.

Thus, the “experiment” lasted only for a year and a half. This is clearly a short time period and implies that the sorts of impacts that could be evaluated under Progresa include only variables likely to be affected in the short run, such as school attendance, expenditure patterns, health clinic attendance, and so on. The evaluation of longer-term variables, for instance, on actual years of completed schooling and future employment of Progresa beneficiaries, marriage and fertility patterns (variables not included in the IFPRI evaluation), will be more difficult to analyze given this short period. Future evaluations of the rural evaluation population using an experimental methodology will only be able to carry out the experiment of comparing, say, effects on families/individuals who received benefits for six years versus the control group of four and a half years, clearly a much weaker comparison than between a family receiving Progresa for six years versus a family who had never received Progresa. For the evaluation of these longer-term variables, some non-experimental methods will also be necessary. This is an important limitation, which we return to at the end of the article.

Constructing a control group is clearly an ethical issue in the context of a program aimed at the poorest members of the population. The point that it might be perceived that eligible communities were being kept out of Progresa for the sake of the evaluation was of great concern to program officials. The argument internally put forth by Progresa officials was that the evaluation was carried out at an early phase where budget limitations implied that not all eligible families could be incorporated. Thus, there were many eligible households who were not receiving benefits during this time period; the actual control group would be no different from these other eligible households; they were simply being monitored (e.g., applied
questionnaires) meanwhile. Indeed, one could argue that in a situation of insufficient funds, the fairest determinant of which families received benefits first would be through a lottery or randomization.

While this argument would seem to be fairly sensible, program officials did not publicize the design of the evaluation until the program evaluation had been concluded in the year 2000 and results were made public. At this point, however, when the evaluation results were released, one prominent newspaper, Reforma, included a number of articles in December 2000 (Torre 2000) criticizing the experimental nature of the evaluation and, in particular, arguing that poor families had been kept out only for the sake of evaluation. The cost of the evaluation (e.g., the contract given to IFPRI) was also criticized, under the argument that the cost of the contract (2.5 million dollars) could have been used to incorporate additional families into the program.

**Information gathering in the evaluation**

In this section, we briefly describe the information that Progresa collected for both the treatment and control group and comment on some issues with the fieldwork.

The first information that Progresa has for both the control group and the treatment group is the survey that is used to identify beneficiary households (Survey of Household Socio-Economic Characteristics [ENCASEH]). For the treatment and control group, it was carried out in October and November 1997. The baseline round of the Evaluation Questionnaire (ENCEL) was carried out in March 1998, approximately one month before the treatment group was incorporated. Follow-up surveys were carried out in November 1998, May 1999, and November 1999. The ENCEL includes extensive information on numerous individual, household, and community characteristics including all sources of income levels, labor market participation, demographic and socioeconomic information, children's school attendance, health utilization, community characteristics, women's status indicators, consumption, asset ownership, and agricultural participation, among others (see Skoufias [2001] for a more detailed description of the data in the different rounds of analysis). Some special interest modules were also carried out, usually on a one-time-only basis, and include such topics as time allocation among different activities and family background of women to measure their bargaining power within the household.

Within each community, all households, regardless of poverty status, were interviewed. This was done partly for evaluation purposes and partly due to the fact that most of the cost of interviewing in communities lies in actually getting to the community. Once there, the marginal cost of carrying out extra interviews is relatively low. This is another important feature of the evaluation design, and it implies that within each treatment community, data exists for both beneficiary and non-beneficiary households; and within each control community, there are data for households who are eligible for Progresa (that is, in extreme poverty) and those who are not eligible for Progresa (Table 5). This is an interesting feature of the data,
which allows one to verify that nonbeneficiary households living in Progresa communities were not subject to spillover aspects (for instance, if they were less likely to receive health services as a result of congestion from Progresa). The nonbeneficiaries can also potentially serve as a control group, particularly in the period after which the control group is incorporated to receive benefits. Nevertheless, there were few evaluation studies in the IFPRI evaluation that actually used the nonbeneficiaries as a control group.

There were a number of issues that Progresa faced in terms of carrying out repeated interviews of the control group. First, the period of the experiment covered a period of huge growth in the number of beneficiaries. Whereas the number of beneficiaries was less than four hundred thousand at the time the experimental period began, there were 2 million households by the time that the experiment was ended. This growth in beneficiaries likely did not go unrecognized by the control communities. Particularly as localities close by (and with quite similar characteristics) began to receive program benefits, interviewers reported more and more complaints within control communities expressed to interviewers there while carrying out follow-up interviews of the ENCEL. In spite of efforts of the interviewers to not reveal that the surveys were related to Progresa, it became quite clear that the control communities were certain that the evaluation surveys were related to Progresa. Pressure applied by control communities through local and state government officials very likely, in combination with the fast growth of Progresa, contributed to the decision by program officials to end the experiment earlier than planned (e.g., at the end of 1999 rather than in the year 2000, which was the original intention). The rapid growth of Progresa implied that the argument that maintaining a control group was not unethical given that many other eligible families were not receiving benefits due to budget constraints was no longer sustainable. This situation is perhaps not unusual, so that one possible lesson here is that randomized designs of evaluations in the case of social programs are unlikely to be able to last for a long period of time, particularly when the constructed control group does not receive any program benefits, as opposed to control groups that might receive some other alternative program during the experimental period.
Evaluation Results and Their Impact on Public Policy

In this section, we turn to the results of the evaluation. We describe the extent to which the randomization was effectively done and how these results affected the estimators used to generate impact results. We also comment on the impact at a public policy level that the evaluation has had, both on policy making in Mexico as well as in other countries. Note that we provide only a brief description of the various impact results of Progresa; an excellent summary of the results is contained in Skoufias (2001). Additionally, the actual evaluation reports can all be found at www.ifpri.org.

Quality of experiment and estimators used in the evaluation

The first aspect of relevance to discuss is the quality of the randomization. That is, are any observed differences between the characteristics of individuals between the two groups statistically significant? This subject is the topic of the first evaluation report carried out by IFPRI, done by Jere Behrman and Petra Todd (1999).

Behrman and Todd (1999) compared characteristics in the treatment and control group for a wide variety of indicators prior to program implementation. In general, they concluded that at the community level, treatment and control groups appear to be random. Nevertheless, at the individual level, they found some small significant differences in preprogram characteristics between the treatment and control group. Behrman and Todd argued that this may largely reflect the large sample size (twenty-four thousand households and more than one hundred thousand individuals).

As a response to the analysis of Behrman and Todd (1999), many of the evaluation reports use the evidence of these differences to justify using regression methods, with control variables, as opposed to simply comparing mean values between the control and treatment group. Additionally, in general, the reports tend to assume that double difference methods are adequate to control for potential preprogram differences in the impact variables of interest, although it is important to note that no report analyzes the extent to which this assumption is justified. Double difference estimators are, in fact, used by almost all of the reports. The only cases in which they are not used are where the relevant impact variables were not included in the baseline. In this case, cross-sectional estimators are generally used to estimate program impacts, that is, comparing differences between the treatment and control group after program implementation. These evaluations, given the evidence of preprogram differences just cited, are the weakest of those carried out in the evaluation.

The standard equation used to estimate double difference impact estimates in almost all of the reports is of the following type:

\[ Y_{it} = \sum_{\tau=1}^{4} \alpha_{\tau} R + \alpha_{T} T_{it} + \alpha_{T} T_{i} + \alpha_{T} T_{i} + \alpha_{T} T_{i} + \beta_{j} X_{j} + \epsilon_{it}, \]
where $Y_t$ reflects the impact variable of interest, $R$ refers to the round of the ENCEL, $T$ refers to whether the individual/household lives in a treatment or control community, and $X_{ijt}$ refers to the vector of $j$ control characteristics for individual/household $i$ in period $t$.

Note that this framework in effect provides double difference estimators of the impact of Progresa. The coefficient $\alpha$ is expected to be statistically insignificant from 0 and provides an indication of whether preprogram differences exist between the treatment group and the control group. The coefficients $\alpha^*_1$ through $\alpha^*_4$ represent the impact estimates, that is, they provide an estimate of the differences between the treatment and control group in the relevant round after program implementation relative to $\alpha$. Note that this framework allows the estimated impacts to differ over time. A final point is that the above equation was generally restricted to the sample of households/individuals eligible for Progresa, that is, most analyses excluded the noneligible from the analysis.

Results of the IFPRI evaluation

The results of the IFPRI evaluation can only be characterized as extremely positive. In the words of the Foreword by Coordinator of the Progresa-IFPRI evaluation, Emmanuel Skoufias,

In summary, the results of the evaluation of IFPRI show that after only three years, poor Mexican children living in the rural areas where Progresa operates have increased their school enrollment, have more balanced diets, are receiving more medical attention, and are learning that the future can be very different from the past. (Skoufias and McClafferty 2001, 3)

Important positive impacts were reported in school enrollment, health clinic attendance, nutrition, and expenditures. In the case of education, the largest impacts were reported on children who enter secondary school, where impacts represent a percentage increase of enrollment of more than 20 percent for girls and 10 percent for boys. The research revealed that much of the positive impact on enrollment was due to increasing continuation rates rather than on getting children who were out of school to return. Little effect of Progresa on enrollment in primary school was found; this, however, reflects the point that most (more than 95 percent) children in rural areas, even before Progresa, enroll in primary school; therefore, the scope for improving enrollment at this level was limited (Schultz 2004; Behrman, Sengupta, and Todd 2000b).

Important health and nutrition effects were also reported. Progresa children aged one to five years have a 12 percent lower incidence of illness than non-Progresa children (Gertler 2000). Additionally, data suggest that Progresa has had large impact on increasing child growth and in reducing child stunting. Behrman and Hoddinott (2000) reported an impact of Progresa equivalent to an increase of 16 percent in mean growth rate per year (corresponding to one centimeter) for children who received treatment between twelve and thirty-six months of age.
Finally, Progresa has also had important effects on expenditure patterns: overall median food expenditures were 13 percent higher in Progresa households when compared with control households. Perhaps just as important, this increase was primarily due to higher expenditures on fruit, vegetables, meats, and animal products, suggesting that beneficiaries are eating a healthier diet under Progresa (Hoddinott, Skoufias, and Washburn 2000).

With respect to the continuation of Progresa, the evaluation likely played an important role in ensuring that the program was not eliminated with the change of government.

Perhaps the only important potentially negative impact found by the IFPRI research team relates to the impacts of Progresa on community relations (Adato 2000). Progresa, as described earlier, in rural areas, targets benefits first at the geographical level and then to households within these eligible communities. This implies that within the same community, there are Progresa beneficiaries and nonbeneficiaries. The IFPRI evaluation reports that this distinction within the community has created some tensions between the two groups.

Impacts of the evaluation on public policy

What impact have these results had on public policy in Mexico? First, with respect to the continuation of Progresa, the evaluation likely played an important role in ensuring that the program was not eliminated with the change of government. In the past few governments, it has been common for each administration, even when the political party of the president is the same, to want to distinguish their years with their own particular social program. In the case of Progresa, whereas there was some ambiguity in the beginning of the Fox administration (the first, by the way, non-PRI (International Revolutionary Party) president in more than seventy years) on Progresa, ultimately, Progresa was not only maintained but expanded to urban areas. In addition, the education grants it provides were extended to the high school level. Note that this particular policy extension was directly recommended in one of the IFPRI documents summarizing the evaluation results. (Skoufias and McClafferty 2001). Thus, the results of the IFPRI evaluation have had an important impact on the direction of the evaluation. This is not to say, however, that Progresa has followed all of the recommendations made by
IFPRI. For instance, the same summary report recommends considering eliminating primary grants and using the money to increase grant levels at the secondary level (where program impacts are much higher). This particular recommendation has not been followed.

With respect to the design of overall social policy in Mexico, the IFPRI evaluation has also had an impact. As mentioned earlier, the IFPRI evaluation was the first rigorous evaluation of a social policy in Mexico. In particular, the year after the IFPRI evaluation was released to the public, Congress issued a new law requiring all social programs to carry out external evaluations of their impacts every year. It is also the case that in the language of this new law, Congress stated that preferably “national institutions” be hired to do the external evaluations, which may have also been a reaction to the IFPRI evaluation.

Finally, at the international level, there have also been important effects of the evaluation. In particular, a number of Latin American countries, such as Colombia, Jamaica, Honduras, and Argentina, have not only implemented Progresa-like programs; they have also implemented the model of external evaluation, in many cases with financing by institutions such as the World Bank and the Inter-American Development Bank.

**Limitations of the Evaluation and Areas for Future Research**

*The sources of the impacts of Progresa*

One serious limitation of the evaluation is that all beneficiaries are offered the same package of benefits. It would be useful to estimate the impact of variation in grant amounts, such as equalizing grants for boys and girls or increasing benefits at the secondary level while removing entire grants from elementary level. Another important question is the effect that the conditionality requirements might have on program impacts, versus a similar program with no conditionality requirements. Moreover, an integrated program of benefits such as the one offered by Progresa makes it difficult to isolate the effect of the different components on relevant outcomes. If an increase in school attendance is witnessed as a result of Progresa, it is of interest to know whether the increase was due entirely to the educational grants received or whether there was also an impact from the health component (because healthier children are more likely to attend school). The design of the evaluation makes it difficult to answer this question (see, however, Rubalcava, Teruel, and Thomas [2004] for an effort to isolate the impact of the mother rather than the father collecting the grant).

A potential, exciting alternative for predicting the impact of potential changes in program design is the use of structural estimation. Todd and Wolpin (2004) estimated a behavior model of parental decisions of child schooling using the Progresa data. They were able to validate the model’s predictions by seeing how well the
model predicts the experimental impact of the program. They then simulated policy interventions that were not part of the randomized design, for instance, varying subsidy amounts. This would seem a promising approach, especially given that even when experimental designs are possible, as in the current case, it is generally infeasible to vary program benefits in many different dimensions to evaluate alternative designs of the program.

**Take-up of benefits**

Another issue that has forth not been studied is that of non-take-up of benefits. Non-take-up here can be broadly defined to include families who do not participate in the program; an alternative indicator would be to focus on children who are eligible for grants but choose not to receive them. Non-take-up overall of families appeared in the evaluation to be initially low at about 3 to 4 percent; nevertheless, for these families, there is no information on reasons for failure to take up the benefits, for instance, whether they moved out of the community or simply refused the program benefits.

Perhaps an even more critical issue is that of non-take-up of children for the education grants in beneficiary families. Whereas there are substantial positive impacts of Progresa on secondary enrollment, there still remain 40 percent of children who are eligible for grants but choose not to receive them. The reasons for this nonparticipation are clearly complex and have not, to our knowledge, been studied. One potential reason is that the size of the education grants is insufficient to motivate the return to school of youth who had potentially already dropped out of school prior to Progresa and thus would have to reenroll in a classroom with much younger classmates.

**Attrition**

Fieldwork protocols in Progresa dictated revisiting only the original dwellings, proceeding to interview whoever was there, if anyone. As a consequence, attrition has turned out to be a serious problem in the ENCEL surveys. For example, by the end of the November 2000 ENCEL round, approximately 16.01 percent of households and 21.89 percent of individuals originally interviewed in the fall of 1997 were no longer in the survey. Most of this attrition is caused by apparent changes of residence or migration (close to 80 percent), and the rest is related to nonresponse and deaths (Teruel and Rubalcava 2003).

Thomas, Frankenberg, and Smith (2001) have highlighted the point that movers may be very different than stayers in the context of longitudinal surveys and stressed the importance of following movers. For the current evaluation, there are two related issues. First, if attrition in the ENCEL surveys differs between treatment and control groups as Teruel and Rubalcava (2003) suggested, this attrition may bias the estimated impacts of the program, and thus it would be useful to have information on the attritors to correct any bias. Second, with high percentages of migrant individuals, it is obviously of interest to know the impacts of the program.
not only on the population remaining in their home communities but also for the
population that leaves. This is likely to be particularly true for youth, who are the
most likely to leave their household and community of origin, and also presumably
the group most likely to be impacted by the program benefits.

In the IFPRI evaluation, none of the evaluation studies considered the possible
biasing effects of attrition on estimated program impacts. Note that following and
locating movers is no trivial exercise, particularly in the Mexican case, where a fair
proportion of movers migrate to the United States. There are now some tentative
plans for the program to follow up a subsample of movers in 2005 or 2006; thus far,
a pilot project has been carried out by Parker (2005) in the state of Queretaro on
following youth aged fifteen to twenty-one who have left their household. The pilot
has been successful at interviewing 70 percent of those moving to a nearby area in
the same state and about 55 percent of those moving to the United States. The
information should be useful for addressing possible program biases and allow
longer-term impacts to be analyzed.

**Outlook**

Perhaps the most important and lasting limitation of the evaluation relates to
the duration of the experimental design, which as described earlier, lasted only
about 1.5 years. While this was adequate for the purposes of evaluating short-term
indicators such as school enrollment or household expenditure—indeed, the eval-
uation was by all indicators quite successful—the longer-term perspective is more
complicated. Given the interest in estimating medium-term impacts of the pro-
gram, a set of new communities that had previously not been eligible for benefits
(because of not satisfying program requirements on having schools and health clin-
ics close by) has been chosen to constitute a new comparison group. A new follow-
up round in 2003 has been carried out, following up the original 506 ENCEL com-
nunities as well as interviewing those in the new comparison communities. In
addition to comparisons based on the experimental design, which are generally
restricted to estimating impacts with respect to differential exposure to benefits (6
versus 4.5 years of benefits), two alternative estimators are now being assessed for
potentially estimating program impact. These include both regression discontinu-
ity analysis with the original ENCEL households and matching between the
ENCEL households and the new comparison group. Whereas the experimental
design has been a critical factor to the program’s initial success, it appears that esti-
mation of the longer-term impacts of the program will need some nonexperimental
estimators or structural estimation such as that carried out in Todd and Wolpin
(2004).

**Notes**

1. The name of Progresa was changed to Oportunidades by President Fox; in this article, nevertheless, we
continue to refer to the program as Progresa.
2. Given this, the rationale for maintaining higher grant levels for girls than boys would seem to be suspect. It is interesting to note that the feature of higher grants for girls than boys was maintained when the program was extended in 2001 to urban areas. This is in spite of the fact that girls in urban areas are currently doing slightly better in terms of overall achievement than boys (Parker and Pedrizzini 2001). In fact, in some Latin American countries, there is a concern of a potential “reverse” gender gap in education (Duryea and Arends-Kuenning 1999), which may also be relevant for Mexico.

3. The states include Guerrero, Hidalgo, Michoacan, Puebla, Queretaro, San Luis Potosi, and Veracruz. Note that in effect there are two baseline surveys that can be used as part of the evaluation, the Survey of Household Socio-Economic Characteristics (ENCASEH) and the first round of the Evaluation Questionnaire (ENCEL).

4. Additional rounds of the ENCEL were carried out in May and November 2000. Both these rounds took place after the control group began to receive benefits and were not used in the evaluation carried out by the International Food Policy Research Institute (IFPRI).

References


Gomez de Leon, Jose, Susan W. Parker, and Daniel Hernandez. 1999. The design and methodology of the impact evaluation of the Education, Health and Nutrition Program (PROGRESA) in Mexico. Photocopy.


