Use of Survey Design for the Evaluation of Social Programs: The PNAD and PETI.

Donald M. Pianto
UFPE

Sergei Soares
The World Bank

July 15, 2004

Área ANPEC: 6 Economia Social

Abstract

The structure of some household surveys allows the evaluation of social programs which are implemented gradually by municipality and whose objectives are measurable by survey variables. Such evaluations do not require over sampling of areas in which the program was implemented, nor the application of additional questionnaires, while providing baseline data and non-experimental comparison groups. We use the PNAD survey to evaluate the impact of the Program for the Eradication of Child Labor on child labor, schooling, and income for municipalities which entered the program from 1997–1999. We present results both from a reflexive comparison and from matching municipalities to form a comparison group and measuring the difference in differences (D in D). Only the reduction of child labor is robust to the D in D analysis, while the reflexive results also demonstrate a significant increase in school attendance. We find the program to be more effective in smaller municipalities as suggested by Rocha (1999).
Keywords: Brazil, Child Labor, Impact evaluation, Propensity score matching
JEL Codes: I32, I38

Resumo

A estrutura de algumas pesquisas por amostragem facilita a avaliação de programas sociais que são gradualmente implementados por município e cujos objetivos são mensuráveis por variáveis da pesquisa. Tais avaliações não exigem a sobre-amostragem de áreas em que o programa foi implementado, nem a aplicação de questionários adicionais, enquanto fornecem dados de *baseline* e grupos de comparação não-experimental. Usamos a PNAD para avaliar o impacto do PETI no trabalho infantil, na escolaridade e na renda dos municípios que entraram no programa entre 1997 e 1999. Apresentamos resultados de uma comparação com a *baseline* e também de uma comparação com um grupo de municípios formado por *matching*. No último caso, medimos a diferença em diferenças (D em D). Apenas a redução do trabalho infantil é robusta à análise D em D, enquanto os resultados de comparação só com a *baseline* também demonstram um aumento significativo na freqüência escolar. Concluímos que o programa é mais efetivo em municípios menores como sugerido por Rocha (1999).
Palavras-chave: Brasil, Trabalho infantil, Avaliação de impacto, *Propensity score matching*
Classificação JEL: I32, I38

1 Introduction

Any social policy or government program should be evaluated. Spending public resources without knowing how they are being used, and what their impact is, is very poor public policy. Process evaluations tell us how these resources were spent and are usually performed closely with day to day monitoring of the program's execution. Without them there is no way to identify and correct blunders in program management. Process evaluations do not, however, tell us what good the program did. This is the territory of impact evaluations.

The ideal social program is designed and implemented with an evaluation strategy that allows for a baseline, a control group, and clear, quantifiable program goals (Baker, 2000; Ravallion, 2001). In the real world this rarely happens. More often than not, political imperatives and the urgency in attacking social problems do not allow for time and resources to be spent designing and implementing an ideal evaluation strategy. The costs involved are usually quite large and well intentioned program directors frequently prefer to spend their money directly on the program's beneficiaries rather than on designing and implementing questionnaires to evaluate their programs. Furthermore, some of the components involved, such as control groups, are often unacceptable from ethical and political points of view. It is difficult to justify that part of a generation be condemned to a lack-luster future to create a control group for program evaluation.

However, most countries have readily available a statistical information system (*e.g.* a household survey) that is collected with other objectives and may be used to evaluate social programs. An evaluation based on household surveys will be performed on a more aggregate level than an evaluation specifically designed for the project. However, the need for evaluations is great and the information available in the household surveys is sufficient to measure program impacts.

The objective of this article is to use Brazilian household surveys to evaluate the impact of a social program that was not designed with evaluation in mind. This is done as a proof of concept that robust evaluations with baseline data can be performed using only household survey data. The program is the Program for Eradication of Child Labor (*Programa de Erradicação do Trabalho Infantil* — PETI) and the survey is the Brazilian National Household Sample Survey (*Pesquisa Nacional por Amostragem de Domicílios* — PNAD).

A favorable World Bank (2001) evaluation of PETI used qualitative methods as well as a quantitative study of twelve municipalities (six with PETI and six without) for which special data was collected. Cardoso and Souza (2003) evaluate the effect of conditional cash transfers on child labor and school attendance in Brazil by using the 2000 census data. They find that transfers increase school attendance while having no effect on child labor. However, as pointed out in Rawlings and Rubio (2003) these studies as well as all other evaluations of PETI lack baseline data. Hence, they are subject to uncorrectable selection biases.

The paper is organized as follows. Section 2 explains how the sampling base for the PNAD is constructed, describes the necessary panel-like characteristic which it contains, and describes the PETI data which was used. Section 3 describes the propensity score matching methodology, its possible biases, and how to correct for them. Section 4 presents the results from our impact evaluation and Section 5 concludes.

2 Data

The PNAD is a cross-section of households throughout Brazil. To reduce costs, the households which are surveyed in the PNAD are not chosen randomly from all those in Brazil. Instead, they are chosen based on a three stage procedure.

The first stage is the choice of municipalities to be surveyed. This is accomplished by first dividing all the municipalities in Brazil (more than 5000) into three categories: (i) metropolitan areas, (ii) auto-representative, and (iii) non-auto-representative. The first two consist of large and medium sized municipalities which are automatically included in the PNAD with probability one. The third category (to which most belong) contains the smaller municipalities which are sampled with probability proportional to their population. The classification and choice of which municipalities are to be surveyed is performed only once during the period between censuses.

During the second stage, the census sectors which will be surveyed are chosen from within the municipalities which have already been selected. These sectors are also maintained throughout the period between censuses. However, areas experiencing rapid population growth will often generate new census sectors which may be included in the PNAD.

In the third stage, the households to be surveyed are chosen from within each census sector. Each year, the households which were interviewed are removed from the PNAD's register and new households are selected for surveying. Therefore, no household is ever surveyed more than once.

Hence, the PNAD consists of a panel of about 800 municipalities, which is fixed between censuses, and a cross-section of households within these municipalities, which changes yearly. This allows one to estimate the impact of a social program on the variables measured by the PNAD when the program is implemented gradually at the level of municipalities.

To complement the PNAD household data, we have data from the Ministry of Social Welfare (*Ministério da Assistência Social* — MAS) detailing which municipalities received PETI and their year of entrance. It is unknown which households from the PNAD participate in PETI, so the only measurable impacts are aggregate ones on municipalities. However, if a PETI child stops working only to have a non-PETI child assume his job, PETI has not been successful. Hence, the restriction to only measure aggregate municipal effects avoids the possibility of positively evaluating PETI in a municipality where child labor didn't decrease, but those children in PETI stopped working.

3 Propensity Score Matching

We can measure part of the impact of PETI by reflexively measuring the changes in the variables of interest in PETI municipalities.[1] Such a measurement describes the changes, but does not attribute them to the program. In order to exclusively attribute any impacts to PETI, we must establish either an experimental control group or a non-experimental comparison group against which we can measure the impact of the program. In the case of PETI, MAS did not randomly choose in which municipalities to implement the program from a larger set of municipalities targeted to receive PETI, hence we cannot pursue an experimental approach. We can,

[1] A reflexive measurement compares the values at a given time after program implementation with the values at a previous time before implementation (a baseline).

however, generate a non-experimental comparison group by propensity score matching (PSM) PNAD municipalities in PETI with other PNAD municipalities which are not in PETI.

PSM techniques are often used on data which do not possess a baseline. In such cases it is necessary to find instruments which affect the probability of participation in a given program, but do not affect the variables of interest, which the program intends to improve. These instruments are then used to match observations (municipalities in our case) which receive the program (the treatment group) with observations that do not receive the program to form a comparison group. Any differences between the variables of interest in the treatment group and the comparison group are attributed to the program impact. This is a typical matching procedure which uses just one cross-section of data.

However, the PNAD provides a baseline for PETI since it is performed annually. Hence, we can match municipalities the year before they enter PETI on instruments which affect the probability of participation *and* on the variables of interest. This is important in the case of PETI, because the criteria by which municipalities are chosen are identical with the variables of interest.

Closely following Heckman, Ichimura, Smith, and Todd (1998) we assume that each municipality has two possible outcomes for a given variable of interest, Y_0 and Y_1, where the first refers to the absence of PETI and the second to its presence. If we define a dummy variable, D, which takes the value one when a municipality receives PETI and zero otherwise, we only observe $Y = Y_1 D + Y_0(1 - D)$ for each municipality.

We would like to measure the impact of PETI, $\Delta = Y_1 - Y_0$. Estimation of this expression using unconditional averages is not recommended since we only observe Y_0 for municipalities that do not receive PETI and this may not be a good estimate of Y_0 for municipalities that do receive PETI.

We can try to observe the mean impact of PETI conditional on characteristics of the municipality, X, given by

$$\Delta(X) = E(\Delta|X, D = 1) = E(Y_1|X, D = 1) - E(Y_0|X, D = 1). \tag{1}$$

We can use the data for municipalities which participate in PETI to calculate $E(Y_1|X, D = 1)$, however $E(Y_0|X, D = 1)$ remains inaccessible. Generally, the equivalence of $E(Y_0|X, D = 1)$ and $E(Y_0|X, D = 0)$ is assumed to allow the estimation of

$$\hat{\Delta}(X) = E(\hat{\Delta}|X) = E(Y_1|X, D = 1) - E(Y_0|X, D = 0). \tag{2}$$

If the number of distinct values of all the variables in X is very large, it becomes difficult to find municipalites with characteristics equivalent to those in the program in order to estimate Y_0 conditional on X. However, PSM relies on a theorem of Rosenbaum and Rubin (1983) to replace the assumed equivalence $E(Y_0|X, D = 1) = E(Y_0|X, D = 0)$ by $E(Y_0|P(X), D = 1) = E(Y_0|P(X), D = 0)$, where $P(X)$ is the probability of participation in PETI. This theorem allows us to replace X by $P(X)$ in Eq. (2) to obtain

$$\hat{\Delta}(P(X)) = E(\hat{\Delta}|P(X)) = E(Y_1|P(X), D = 1) - E(Y_0|P(X), D = 0), \tag{3}$$

where $P(X)$ is estimated by a probit model. Matching is now performed on a single number.

Our interest is the averaged version of this equation (Heckman, Ichimura, Smith, and Todd, 1998, eq. 2)

$$\hat{\Delta}(K) = \int_K \hat{\Delta}(P(X))\mathrm{d}F(P(X)|D = 1) \Big/ \int_K \mathrm{d}F(P(X)|D = 1), \tag{4}$$

where K represents the region in probability-space over which the impact is being averaged, and the argument K on the left hand side implies averaging over this region.

The resulting estimate from matching on the propensity score $\hat{\Delta}(P(X))$ contains a bias,

$$B(P(X)) = E(Y_0|P(X), D = 1) - E(Y_0|P(X), D = 0),$$

such that $\hat{\Delta}(P(X)) = \Delta(P(X)) + B(X)$. The average of this bias term is decomposed into three components (Heckman, Ichimura, Smith, and Todd, 1998, eq. 14, p. 1030). The first term arises if we calculate $\hat{\Delta}(K)$ in regions K where there is no common support for PETI and non-PETI municipalities. This means that a given municipality may not have a match with similar characteristics and therefore should not be used in the estimate of program impact. We can eliminate this bias by requiring that all matches be made within a limited radius of the PETI municipalities' propensity scores as in the "propensity-score caliper" of Dehejia and Wahba (1998) used by Jalan and Ravallion (2003):

$$C[P(X_i)] = \{P(X_j) \mid \|P(X_i) - P(X_j)\| < \epsilon\} \tag{5}$$

Municipalities which do not manage a match are not used in the estimate of $\hat{\Delta}(K)$ in Eq (4).

The second component of the bias term is generated by the difference in the density of PETI and non-PETI municipalities as a function of the propensity score. We eliminate this bias by re-weighting the density of non-PETI municipalities through matching. We obtain Y_1 for each PETI municipality that is matched to one or more non-PETI municipalities. We then calculate an average value of Y_0 for the non-PETI municipalities matched to that municipality, finally calculating (Jalan and Ravallion, 2003)

$$\hat{\Delta} = \sum_{m=1}^{P} N_m \left(Y_{1m} - \sum_{n=1}^{\text{NonP}} W_{nm} Y_{0n} \right) \Bigg/ \sum_{p=1}^{P} N_p, \tag{6}$$

where P is the number of matched PETI municipalities, Non P is the number of non-PETI municipalities, N_m is the number of observations in municipality m, and W_{nm} defines the re-weighting of the non-PETI municipalities. For nearest neighbor matching W_{nm} equals one for the nearest neigbor and zero for all others. In general, we weight matches by one over the number of matches obtained for municipality m. This re-weighting of the non-PETI municipalities eliminates the second component of bias. The final component of the bias term is not eliminated by matching and is the true selection bias, $B(P(X))$.

Heckman, Ichimura, Smith, and Todd (1998) find that although $B(P(X))$ may not be zero for a given cross-section, if it is constant in time an accurate estimate of $\delta\Delta^{(t)} \equiv \Delta^{(t)} - \Delta^{(t-1)}$, for a municipality which enters PETI in year t is given by the difference in differences (D in D) estimator. Assuming the bias is constant in time

$$\begin{aligned} \delta\Delta^{(t)}(P^{(t)}(X^{(t-1)})) \quad = \quad & \hat{\Delta}^{(t)}(P^{(t)}(X^{(t-1)})) - \Big[E(Y_0^{(t-1)}|P^{(t)}(X^{(t-1)}), D^{(t)} = 1) \\ & - E(Y_0^{(t-1)}|P^{(t)}(X^{(t-1)}), D^{(t)} = 0) \Big], \end{aligned} \tag{7}$$

where: $P^{(t)}(X^{(t-1)})$ represents the probability of a municipality entering PETI in year t based on its characteristics in year $t-1$ and will be referred to as $P^{(t)}$ in the future for brevity; $D^{(t)}$ is one if a municipality enters PETI in year t, zero if it is not in PETI in year t, and missing if it has entered PETI before year t; and $Y_0^{(t-1)}$ represents the value of Y in year $t-1$ for municipalities in which PETI has not been implemented in year $t-1$. The term in brackets, $B^{(t-1)}(P^{(t)})$, is directly calculable from the data in year $t-1$, because PETI has not been implemented yet and all municipalities used in the calcuation have $D^{(t-1)} = 0$ and hence $Y^{(t-1)} = Y_0^{(t-1)}$. Reiterating, this estimate is accurate if $B^{(t)}(P^{(t)}) = B^{(t-1)}(P^{(t)})$.

We rewrite the above D in D equation as

$$\begin{aligned}
\delta\Delta^{(t)}(P^{(t)}) = \quad & \left(E(Y_1^{(t)}|P^{(t)}, D^{(t)} = 1) - E(Y_0^{(t)}|P^{(t)}, D^{(t)} = 0)\right) \\
- \quad & \left(E(Y_0^{(t-1)}|P^{(t)}, D^{(t)} = 1) - E(Y_0^{(t-1)}|P^{(t)}, D^{(t)} = 0)\right) \\
= \quad & \left(E(Y_1^{(t)}|P^{(t)}, D^{(t)} = 1) - E(Y_0^{(t-1)}|P^{(t)}, D^{(t)} = 1)\right) \\
- \quad & \left(E(Y_0^{(t)}|P^{(t)}, D^{(t)} = 0) - E(Y_0^{(t-1)}|P^{(t)}, D^{(t)} = 0)\right) \\
\equiv \quad & \left(\delta Y(D^{(t)} = 1)\right) - \left(\delta Y(D^{(t)} = 0)\right), \qquad (8)
\end{aligned}$$

to show how PETI's impact as measured by D in D can be written as the difference between the reflexive estimate of its impact, $\delta Y(D^{(t)} = 1)$, and the change in the untreated comparison group, $\delta Y(D^{(t)} = 0)$. We report the averaged version of this impact, $\delta\Delta^{(t)}(K)$, obtained by replacing Y_{1m} and Y_{0n} in Eq. 6 by $\delta Y_m(D^{(t)} = 1)$ and $\delta Y_n(D^{(t)} = 0)$, respectively.

4 Results

4.1 The PNAD

Table 1 reports the number of municipalities in all of Brazil and the number of municipalities surveyed by the PNAD that entered PETI from 1997–1999. The data suggest that the municipalities in the PNAD provide good coverage of the municipalities which received PETI.

We must ensure that the PNAD is providing a similar cross-section of households within each municipality from year to year. To this end we perform t-tests for the equality of means between the average adult education the year before entering PETI and the year of entrance. Adult education should not change significantly from one year to the next, hence any differences come from sampling errors. Any municipalities in which the null hypothesis of equal means can be rejected at a significance level of 5% are excluded from the probit and are not used when matching, including PETI municipalities. The numbers of excluded municipalities are 88, 58, and 74 for the years from 1996–7, 1997–8, and 1998–9, respectively. Two PETI municipalities are excluded for the 1997-8 period and none are excluded in the other periods.

The PNAD is a survey which provides sampling weights and information on the primary sampling units (PSUs). When measuring the impact of PETI we use neither the sampling weights nor the PSUs. The use of the weights would be an error since our goal is not to extrapolate the impact of PETI as if it was implemented throughout Brazil, but rather to measure its impact on the specific municipalities in which it was implemented. The use of the PSU information would be acceptable, however, we prefer to treat each observation as independent, rather than the PSUs. Therefore, the estimates are performed using analytic weights (cell-mean weights) where the weights are the number of observations in the municipality, not the number of PSUs (N_m in Eq. 6). This approach causes municipalities with more observations (or PSUs) to have a larger effect on the estimates of PETI's impact. This is troublesome since it has been suggested (Rocha, 1999) that PETI would be more effective if it were implemented on a smaller scale (neighborhoods) rather than a larger one. We resolve this problem by including a dummy variable in our regressions which indicates which PETI municipalities have more than 800 observations in the PNAD.

The main objective of PETI is to reduce child labor in targeted degrading or dangerous activities, as enumerated in appendix II of its orientation guide (MEAS, 2002). These targeted activities include working in coal-pits, construction, any work with dangerous objects or chemicals, and many more. Households with a per-capita income less than one half of the minimum salary are preferentially targeted. In PETI households, the children are required to attend school and after school programs, and to stop working. Further details of PETI's functioning can be found in MEAS (2002) and Rocha (1999).

In order to assess the efficacy of the program several variables measured by the PNAD are used. Child labor for children from 10–14 years old is measured, including the type of labor, which allows the classification into targeted rural and urban activities and non-targeted activities. School attendance is measured for all respondents, although we only measure the effect on children from 7–14 years of age. Years of study is measured for all respondents, allowing the calculation of how far off track children from 7–14 are from their expected educational level, based on age. Household income is reported, which allows the calculation of per-capita household income and the percentage of households with per-capita income less than 25% of the minimum salary.

These variables are used to measure the impact of PETI on child labor, schooling, poverty, and income. The reflexive changes measured in these variables for PETI municipalities are independent of the probit and matching results given below and are reported in columns 2, 4, and 6 of Tables 3 and 4 for municipalities which entered PETI from 1997–1999 and those which entered in 1998, respectively.

4.2 Matching

As discussed in Section 3, since our final goal is a D in D measurement, we can match based on the same variables on which we would like to measure the impact of PETI. We present two different probit models which are used to determine the probability of a given municipality entering PETI.

The first probit specification in the left panel of Table 2 is poorly adjusted, but jointly significant to 3%. The variables used in the probit represent average municipal values and are: per-capita income; Theil's T measure of income inequality; average adult education in years for adults 25 and older; percentage of illiterate adults 25 and older; percentage of children 7–14 who are more than one year behind in school; percentage of children 10–14 who work in urban and rural activities targeted by PETI and who work in untargeted activities; percentage of people who live in urban as opposed to rural areas; and the type of census zone to which the municipality belongs. These variables are used in the hope that a comparison group can be formed with the same poverty, inequality, and educational deficiencies, as well as similar child labor levels, as the municipalities that receive PETI.

Three variables in the left panel of Table 2 are significant. The coefficient of income demonstrates that the higher the per-capita income is the less likely a municipality is to enter PETI. This is to be expected if we assume that children are more likely to work if their family needs their earnings. The large positive coefficient of not attending school shows that the smaller the percentage of children going to school, the higher the probability of entering PETI. Again, this is an expected result since children who work are less likely to attend school. And finally, being in a non-auto-representative municipality significantly decreases the probability of entering PETI. We used many combinations of child labor variables in an attempt to find a significant coefficient for child labor, but found none.

Although the model is poorly adjusted, one can see in columns 1 through 3 in Table 3[2] that the comparison group formed by a simple nearest neighbor PSM using this probit specification reasonably reproduces the PETI treatment group. Income and the high levels of child labor could not be reproduced in the comparison group, which is not surprising since none of the child labor variables were significant in the probit which we used.

The second probit specification in the right panel of Table 2 is better adjusted than the first and is jointly significant to less than 0.1%. The specification is similar to the first, with the exception of two variables: off track now represents the average number of years that a child 7-14 is behind in school, rather than a dummy variable; and the child labor variables have been replaced with adult labor in activities targeted by PETI.

This model also has only three variables with significant coefficients. The coefficients of non-auto-representative and not attending school exhibit the same behavior as in the first model. The coefficient of the new variable, adult labor in targeted activities, is significant to 1% and demonstrates that the higher the incidence of adult labor in jobs considered degrading for children, the greater the likelihood that a municipality will enter PETI. This indicates that MAS did not necessarily target municipalities with high incidences of child labor in degrading activities, but rather municipalities where there was a high risk of children being employed in dangerous activities because of high demand for such labor.

Looking at columns 1 through 3 in Table 4 one sees that the comparison group formed by PSM with this probit is excellent. The PSM was performed with up to four nearest-neighbors with propensity scores within a radius of 0.005 of the PETI municipality's propensity score. Three matchable PETI municipalities failed to find matches because of this common support restriction. Additionally, two other PETI municipalities had more than 800 observations in the PNAD and are not represented in the table.

4.3 Difference in Differences

The sixth and last columns of Tables 3 and 4 contain results pertaining to a purely reflexive estimate of the impact of PETI and the D in D estimate, respectively. In both tables, column 6 minus column 7 gives column 8 as demonstrated in Eq. (8).

Table 3 contains the results from aggregating all municipalities that entered PETI from 1997–1999. The variables on which the impact of PETI is measured are described in Section 4.1. There are five statistically significant results from the reflexive estimate and two statistically significant results from the D in D estimate. We only comment on those results which are significant to at least 5%. Both estimates show a decrease in all child labor of approximately 8 percentage points which is significant to 1% and a decrease in child labor in targeted areas of about 4 percentage points, significant to 5%. The reflexive results also demonstrate an increase

[2]The same probit specification was used as that in the left panel of Table 2 with data for municipalities that entered PETI from 1997–9. There are 2092 observations in column 1 because each municipality receives a different propensity score and contributes different data depending on which year the probit is being run.

in school attendance of about 4 percentage points, significant to 1%. Although not reported in the table, the four PETI municipalities which had more than 800 observations in the PNAD showed no significant PETI impacts, neither for the reflexive nor the D in D estimates.

Table 4 contains the results of the impact evaluation for the municipalities which entered PETI in 1998. The reflexive results show a significant increase in per-capita income, a 6.9 percentage point increase in school attendance, significant to 1%, and a decrease of about 10 percentage points from 22 to 12% in all child labor, significant to 5%. The D in D results show a 10.8 percentage point decrease in all child labor, significant to 1%, as well as a 4 percentage point decrease in targeted child labor, significant to 10%, both attributable to PETI.

Although not reported, the two PETI municipalities which had more than 800 observations in the PNAD in 1997 had an increase in per-capita income of R$15, significant to 10%, based on the reflexive estimate. However, the D in D estimate yielded many disturbing results: the percentage of people living with a per-capita income less than one quarter of the minimum salary rose 8.4 percentage points; the average number of years that a child is behind in school rose 0.25; and the percentage of children who attend school fell 5.2 percentage points; all significant to 5%. Rocha (1999) suggests that PETI be targeted not at the municipal level, but at the neighborhood level to ensure proper targeting and delivery of benefits. Our results demonstrate that PETI is more effective in smaller municipalities, perhaps for the reasons suggested by Rocha (1999).

The two matching models reported in this paper have a large overlap in the data which is used to measure the impact of PETI, since 60% of the PETI municipalities which entered the program from 1997–1999 entered in 1998. This explains the similarity of the reflexive results which show increases in school attendance and decreases in child labor. However, we use different probit models and matching criteria on the different data sets. Despite these differences the models agree that PETI has significantly decreased child labor in the municipalities in which it was implemented as measured by a D in D estimate. These general results are robust to variations in the number of nearest-neighbors matched from 1–5 (and maybe more) and variations in the radius of matching from 0.005–0.025 for the municipalities which entered PETI in 1998 and are available from the authors upon request.

5 Conclusion

We have exploited the properties of the PNAD to perform an impact evaluation of the Program for the Eradication of Child Labor (PETI) in Brazil. Our results demonstrate that PETI works better in smaller municipalities in line with suggestions by Rocha (1999) that the program be implemented on smaller scales to improve its effectiveness. PETI has a significant impact on the reduction of child labor as measured by a difference in differences (D in D) estimate. Although a reflexive estimate shows significant increases in school attendance, these results do not survive the D in D analysis. This may be a result of our lack of knowledge of what other social programs have been implemented in the various municipalities. For example, *Bolsa Escola*, which provides monthly stipends for children to attend school, but has no requirement that these children not work, could be masking the effect of PETI on school attendance if the municipalities in the comparison group are receiving this program. In order to correct such a problem it would be necessary to gather data about which municipalities receive which social programs. Then these municipalities could be kept out of the comparison group, or a more complex matching could be performed which measures the relative impacts of different programs.

The structure of the PNAD was particularly well suited to measuring the impact of PETI, since the most important effects of the program (reducing child labor and increasing school attendance) only have significance on a municipality wide scale. Hence, the aggregation of the data was not detrimental to evaluating PETI. However, an oversampling of the municipalities which receive PETI, with an additional oversampling of PETI participants within those municipalities, and a question identifying which households receive PETI could be extremely useful in future evaluations of PETI (Jalan and Ravallion, 2003). First, it might become possible to measure impacts on poverty and income for PETI households. Second, the information could be used to monitor the targeting of PETI within the municipalities. Also, the grouped estimator of Dehejia (2003) could be used to predict which municipalities might respond better to PETI even before implementation. However, it is important to note that we have managed to perform an evaluation without the need for oversampling or any extra costs or questions beyond those which already exist in the PNAD.

In Brazil, our methodology can be used to evaluate any social program: which is implemented gradually on the scale of municipalities; which intends to affect variables measured by the PNAD; and whose impact is only effective on a community-wide scale. Generally speaking, any household survey with a panel characteristic similar to that of the PNAD and with a sufficient number of variables of interest can be used to perform a similar impact evaluation.

References

BAKER, J. L. (2000): *Evaluating the Impact of Development Projects on Poverty A Handbook for Practitioners*. The World Bank, Washington, D.C.

CARDOSO, E., AND A. P. SOUZA (2003): "The Impact of Cash Transfers on Child Labor and School Attendance in Brazil," mimeo, Universidade de São Paulo.

DEHEJIA, R. H. (2003): "Was There a Riverside Miracle? A Hierarchical Framework for Evaluating Programs With Grouped Data," *Journal of Business & Economic Statistics*, 21, 1–11.

DEHEJIA, R. H., AND S. WAHBA (1998): "Propensity Score Mathcing Methods for Non-experimental Causal Studies," Working Paper 6829, NBER.

HECKMAN, J., H. ICHIMURA, J. SMITH, AND P. TODD (1998): "Characterizing Selection Bias Using Experimental Data," *Econometrica*, 66(5), 1017–1098.

JALAN, J., AND M. RAVALLION (2003): "Estimating the benefit incidence of an antipoverty program by propensity-score matching," *Journal of Business & Economic Statistics*, 21(1), 19–30.

MEAS (2002): *Programa de Erradicação do Trabalho Infantil — PETI, Manual de Orientações*Ministério da Previdcia e Assistência Social, Brasília, DF, Brasil, 1 edn.

RAVALLION, M. (2001): "The Mystery of the Vanishing Benefits: An Introduction to Impact Evaluation," *World Bank Economic Review*, 15(1), 115–140.

RAWLINGS, L. B., AND G. M. RUBIO (2003): "Evaluating the Impact of Conditional Cash Transfer Programs: Lessons from Latin America," Working Paper 3119, World Bank.

ROCHA, S. (1999): "Child Labor in Brazil and the Program for Eradication of Child Labor A case study of Pernambuco," Background paper, World Bank.

ROSENBAUM, P., AND D. RUBIN (1983): "The Central Role of Propensity Score in Observational Studies for causal Effects," *Biometrika*, 70.

WORLD BANK (2001): "Brazil: Eradicating Child Labor in Brazil," Report No. 21858-BR, World Bank.

Table 1: Number of Municipalities that entered PETI by year.

	All municipalities		PNAD municipalities	
Year	Number	Percent	Number	Percent
1997	33	21.57	7	17.50
1998	93	60.78	24	60.00
1999	27	17.65	9	22.50
Total	153	100.00	40	100.00

Table 2: Two probit specifications for municipalities which entered PETI in 1998.

Variable	Coefficient	Std. Error	Variable	Coefficient	Std. Error
Income (R$)	-0.00438	0.00267^{*}	=	-0.00417	0.00262
Theil's T	-0.06819	0.55638	=	0.02479	0.56244
Adult Education	0.07009	0.17079	=	0.09020	0.17541
Adult Illiteracy	1.87471	1.47229	=	1.83464	1.53788
Off Track > 1	-1.40419	1.02196	Off Track	-0.37360	0.28870
No School	3.56533	$1.44443^{\dagger}$	=	3.30303	$1.48183^{\dagger}$
Child Labor			Adult Labor		
Targeted Urban	-1.89193	4.66214	Targeted	3.62463	$1.15713^{\ddagger}$
Targeted Rural	1.42717	1.29894			
Untargeted	0.66858	0.91179			
Urban area	0.69960	0.59431	=	0.08130	0.56738
Census Zone			=		
Metropolitan	—	—	=	—	—
Auto rep.	0.05960	0.34629	=	0.01996	0.34867
Non-auto rep.	-0.56402	0.32165^{*}	=	-0.60963	0.32540^{*}
Intercept	-2.00668	1.03688^{*}	=	-2.62992	$1.06671^{\dagger}$
LR $\chi^2(12)$	23.73		LR $\chi^2(10)$	31.51	
p-value	0.0221		p-value	0.0005	

Source: IBGE/PNAD 1997; MAS/PETI. **Note:** Standard errors in parenthesis. ‡, †, and * represent significance at the level of 1, 5, and 10%, respectively. An '=' in the right 'Variable' column means that the variable is the same as that in the left 'Variable' column.

Table 3: Baseline data and PETI impact for municipalities that entered PETI from 1997–1999 and have 800 or fewer observations in the PNAD.

	Year before entering			Entrance year		δY		PETI
	Non-PETI	PETI	Comp.	PETI	Comp.	PETI	Comp.	Impact
Variable	Mun.	Mun.	Mun.	Mun.	Mun.	Mun.	Mun.	$\delta\Delta(K)$
Income (R$)	231.66 (2.93)	119.76 (10.95)	162.30 (13.85)	126.13 (11.86)	165.09 (13.56)	6.37 (4.21)	2.79 (5.27)	3.57 (7.19)
% Poor	0.110 (0.003)	0.179 (0.022)	0.168 (0.023)	0.205 (0.024)	0.203 (0.022)	0.026 (0.016)	0.035 (0.014)†	−0.009 (0.022)
Off Track	1.185 (0.013)	1.652 (0.090)	1.568 (0.104)	1.556 (0.090)	1.577 (0.102)	−0.096 (0.051)*	0.009 (0.054)	−0.105 (0.076)
Go to School	0.935 (0.001)	0.877 (0.012)	0.911 (0.010)	0.918 (0.009)	0.927 (0.011)	0.041 (0.014)‡	0.016 (0.013)	0.025 (0.019)
Child Labor	0.131 (0.003)	0.263 (0.031)	0.156 (0.022)	0.181 (0.027)	0.157 (0.024)	−0.082 (0.024)‡	0.001 (0.016)	−0.083 (0.028)‡
Targeted	0.039 (0.001)	0.099 (0.016)	0.050 (0.009)	0.066 (0.015)	0.060 (0.014)	−0.034 (0.016)†	0.010 (0.011)	−0.044 (0.019)†
Obs.	$n = 2092$	$n = 34$	$n = 38$	$n = 34$	$n = 38$	$n = 34$	$n = 38$	$n = 38$

Source: IBGE/PNADs 1996–1999; MAS/PETI. **Note:** Standard errors in parenthesis. ‡, †, and $*$ in the last three columns represent significance at the 1, 5, and 10% level, respectively. Propensity score matching was performed by using the probit specification in the left panel of Table 2 estimated for all municipalities which entered PETI from 1997–9. A simple nearest neighbor matching was performed.

Table 4: Baseline data and PETI impact for municipalities that entered PETI in 1998 and have 800 or fewer observations in the PNAD.

Variable	1997 Values Non-PETI Mun.	1997 Values PETI Mun.	1997 Values Comp. Mun.	1998 Values PETI Mun.	1998 Values Comp. Mun.	δY PETI Mun.	δY Comp. Mun.	PETI Impact $\delta\Delta(K)$
Income (R$)	238.78 (5.14)	129.81 (19.49)	121.85 (10.36)	142.32 (20.62)	134.13 (11.86)	12.51 (5.15)†	12.28 (3.55)‡	0.23 (7.47)
% Poor	0.109 (0.005)	0.168 (0.032)	0.185 (0.022)	0.181 (0.031)	0.184 (0.019)	0.014 (0.023)	−0.002 (0.013)	0.015 (0.033)
Off Track	1.195 (0.022)	1.651 (0.158)	1.557 (0.114)	1.497 (0.148)	1.492 (0.119)	−0.154 (0.081)*	−0.065 (0.035)*	−0.089 (0.102)
Go to School	0.936 (0.002)	0.864 (0.021)	0.875 (0.017)	0.932 (0.008)	0.923 (0.012)	0.069 (0.019)‡	0.048 (0.012)‡	0.021 (0.016)
Child Labor	0.131 (0.005)	0.222 (0.046)	0.201 (0.025)	0.118 (0.020)	0.205 (0.032)	−0.105 (0.037)†	0.004 (0.016)	−0.108 (0.034)‡
Targeted	0.038 (0.002)	0.083 (0.025)	0.083 (0.014)	0.044 (0.009)	0.084 (0.015)	−0.039 (0.025)	0.001 (0.010)	−0.040 (0.022)*
Obs.	$n = 707$	$n = 17$	$n = 17$	$n = 17$	$n = 17$	$n = 17$	$n = 17$	$n = 17$

Source: IBGE/PNADs 1997 and 1998; MAS/PETI. **Note:** Standard errors in parenthesis. ‡, †, and ∗ in the last three columns represent significance at the 1, 5, and 10% level, respectively. Propensity score matching was performed by using the probit specified in the right panel of Table 2. The matching was performed with up to four nearest neighbors whose propensity scores were within 0.005 above or below that of the matched PETI municipality. Three PETI municipalities were not matched because of this last restriction.