# Courant Research Centre
## 'Poverty, Equity and Growth in Developing and Transition Countries: Statistical Methods and Empirical Analysis'

**Georg-August-Universität Göttingen**
**(founded in 1737)**

Discussion Papers

**No. 84**

**Conditional Cash Transfers in Brazil: Treatment Evaluation of the "Bolsa Família" Program on Education**

**Elke Schaffland**

**July 2011 (revised April 2012)**

Wilhelm-Weber-Str. 2 · 37073 Goettingen · Germany
Phone: +49-(0)551-3914066 · Fax: +49-(0)551-3914059

Email: crc-peg@uni-goettingen.de Web: http://www.uni-goettingen.de/crc-peg

# Conditional Cash Transfers in Brazil: Treatment Evaluation of the "Bolsa Família" Program on Education

Elke Schaffland[1]

Courant Research Centre „Poverty, Equity and Growth"

Georg-August-University Göttingen, Germany

[1] I thank Stephan Klasen, Jan Priebe, Chris Murris, Rodolfo Stucchi, Jenny Aker, Elisabeth Sadoulet, Allain de Janvry, Cornelia Römling, Nicole Grunewald and Christian Bjornskov for their comments and support, and Rita Motzigkeit and Laura Stapff for excellent research assistance.

## Abstract

Brazil's "Bolsa Família" conditional cash transfer program (BFP) is the most substantial poverty alleviation program in Brazil. It is the biggest program of this kind in the world, reaching more then 13 000 000 families. The BFP awards grants to eligible poor families, allowing increased consumption in the short term, while building human capital in the long term through setting requirements for school attendance and health care. In this paper, we evaluate the effect of these transfers on educational outcomes in 2004 and 2006, as well as heterogeneous treatment effects over age, region, gender and area (rural/urban). Using Propensity Score Matching (PSM), we find that the probability of school enrollment rises by around 4 percentage points for recipients' children. The effect is higher for younger children, as well as children living in less developed regions (north and north east) and rural areas. Our results also point to a positive impact on school attendance; recipients miss less school than non-recipient children. However, we also find that this result slightly fades over time.



Keywords: conditional cash transfers, propensity score matching, education

## Introduction

Brazil is known for its relatively high growth rates during the 80's and 90's, while maintaining relatively poor human development indicators. The persistent inequality in the country and lack of social improvement in the last decade are prevailing factors which explain the high growth rates in the face of high inequality. In an effort to change this situation, policies for income redistribution in Brazil began to be implemented nationwide in the year 2001. One of these policies is the widespread and well-known "Bolsa Família Program" (BFP), a Conditional Cash Transfer (CCT) program which has benefited 12 000 000 families to date. The intentions are twofold. First, in the short-run, the program attempts to increase income and alleviate poverty. Second, in the long-run, it attempts to break the inter-generational cycle of poverty by increasing human capital.

The program gives grants to families which fit the principal inclusion criteria for participation. This means having a per capita income of less than R$140 per month, and therefore, being below the poverty line. As a condition, families need to send their children to school and attend at least 85% of classes, follow the childrens' vaccination calendar and be submitted to prenatal medical care. The BFP is now considered the biggest program of its kind in the world and one of the drivers of decreased inequality in Brazil. Its impact has been estimated as being responsible for a 20 – 25% reduction in inequality, and the recent 16 % decrease in extreme poverty (Barros et al., 2006).

Cash transfer programs have been exhaustively analyzed in Central and Latin America. The PROGRESA experience, a randomized field experiment in Mexico, is the most frequently analyzed cash transfer program. The program has a positive effect on enrollment, dropout, repetition rates and test scores (Behrman, Sengupta and Todd, 2005 and 2000; Schultz, 2004) as well as it properly work as a safety net for education among the poor (de Janvry et al., 2006). Additionally to these findings, CCTs in Nicaragua increased enrollment in 13% (Maluccio and Flores 2005) and improved language and personal behavioral skills (Macours, Schady and Vakis 2008). In Honduras, school enrollment rose by 3 % (Glewwe and Olinto, 2004), in Ecuador by 10 % for recipients in first income quintile (Schady and Araujo, 2008 and Oosterbeek et al. 2008) but had no effect on test scores (Ponce and Bedi, 2008). In Jamaica, Levy and Ohls (2010) find a rise of 0.5 days per month in attendance for recipient children.

All these results show a positive effect on education. Surprisingly, there have been less empirical studies on the Brazilian CCT, probably due to the fact that the program was never implemented using randomized treatment. Bourguignon et al. (2003) provide an ex-ante analysis of the impact of “Bolsa Escola” on school enrollment. By using a micro-econometric simulation the authors estimate that 60% of poor children aged between 10 and 15 enroll to school due to program participation. They find also find a low effect on poverty and inequality, however.

Using PSM, Cardoso and Portela Souza (2004) analyze the impact of the “Bolsa Escola” program utilizing the 2000 census data. Their results show that treated children were 3 to 4 percentage points more likely to go to school. However, due to a small number of covariates used under PSM, the quality of the control group might be lacking, causing a significantly high unobservable bias. Furthermore, the dataset does not allow for the distinction between “Bolsa Escola” and former transfer programs in Brazil.

De Janvry et al. (2006) analyze the impact of “Bolsa Escola” in the northeastern states of Brazil. The authors used data from two dimensions, a survey of 261 municipalities and a compilation of school records from 1999 to 2004, finding an overall reduction of 8 percentage points in the dropout rate of recipients, and an increase of 0.8% in the failure rate. Glewwe and Kassouf (2008) also find a reduction of 8 percentage points in the dropout rate in primary schools. However, due to the use of the cash transfer variable at the municipal level, the analysis is rather weak; the effect of “Bolsa Escola” was only caught by the data if the program was operating at that time in the municipality.

Another contribution to the area was a report solely dedicated to the “Bolsa Família”, launched by the Ministry of Social Development and Fight Against Hunger in 2007. Oliveira et al. (2007) compare the effect of “Bolsa Família” with other cash transfers which are not conditional, finding no positive effect of conditionality on educational output when compared to other cash transfer programs. Yet, when comparing to children with no benefit at all, they find evidence of a positive effect of ”Bolsa Família” on attendance, as well as lower proportion of children dropping out, but a negative effect in regards to passing the school year. However, we raise some concerns regarding the method and the results. Firstly, they use a small set of variables to build the comparable counterfactual group, permitting a higher unobservable bias, while not reporting any results on balancing or common support. Additionally, the authors analyze different regions and gender effects without analyzing if the differences are statistically different.

It is surprising that so few evaluations have been done on this program, and those that have each have methodological or empirical shortcomings. The program is the biggest CCT programs in the world and belongs to the first generation of its kind, commencing in 1997 together with PROGRESA. We innovate by analyzing the effect of "Bolsa Família" on two periods of time, 2004 and 2006, providing an analysis of the lasting impact of "Bolsa Família" on enrollment and attendance, using country wide micro data at the individual level. Due to a rich level of data, we are able to construct a good counterfactual group for the analysis, which we consider to be an improvement in methodology. We are also able to provide an overall impact assessment as well as analyzing heterogeneous treatment effects caused by age, region, gender and area (rural, urban, slums). In this way, we are able to contribute to the broad analysis of the impact of CCTs in a more accurate and encompassing manner. Additionally we find that the positive impact of the program slightly fell by 1% between the years of 2004 and 2006.

The results show that there is a positive impact on educational output driven by program participation. The probability of enrollment rises by around 4 percentage points for recipient children aged between 6 and 17. We also found a positive impact on school attendance, with the same group of children estimated to miss around 0.3 less days of school (during the preceding two months) than non-beneficiary children. We also looked at differences between regions, areas and gender, finding that the program's impact is higher in less developed regions (North and North-East) and rural areas.

## 1.1 The Bolsa Família Program

The Bolsa Família Program gives grants to each child of poor families, as well as a basic grant to families living in extreme poverty (currently with per capita income less then R$70). The grant varies from R$32 to R$306 depending on monthly per capita income and the number of children aged between 0 and 17 living in the household. There are three kinds of benefits which are distributed based on the level of poverty, number of children and children's age. The "basic benefit" assigns R$ 70 to families in extreme poverty independent of the number of children. The "variable benefit" grants poor families R$32 for each child aged between 0 and 15 and the “teenager benefit” grants children between 16 and 17 years old who attend school R$38. [2]

To provide a comparison concerning the per capita amount granted to recipients, Diaz and Handa (2006) state that, to have a significant impact, a rule of thumb regarding the amount of cash transfers is that they remain between 20% and 40% of the per capita income amount which sets the poverty line. Compared to other Latin American countries, in 2006, which is the year of data collection, Brazil was one of the two countries (the other is Honduras) which did not reach this grant amount (Diaz and Handa, 2006). Particularly for Brazil, the poverty line in 2006 was set at R$120, while the grant per child was only R$20, which was 16.6% of the per capita poverty line. However, Brazilian government is frequently raising the grant. Nowadays the grant is 23% of the per capita poverty line.

Once in the program, the family receives a grant for all its children. The participation is conditional on at least 85% school attendance for children between the ages of 6 to 15. For children of ages 16 and 17, the attendance obligation for recipients falls to 75%. Regarding medical care, parents are obligated to provide a basic level of this, which means allowing children to get required vaccines up until they turn 6, as well as prenatal medical assistance and breast feeding mothers.

The cash transfer is granted for the household, and not at the individual level. Therefore, once in a beneficiary household, all children are bounded to the school attendance conditionality. However, a teenager who has already turned 16 has the right to abdicate his grant and therefore does not need to meet the conditionality if he decides to work instead of study.[3]

### 1.1.1 Program structure and targeting scheme

Considering that the program design is clearly based on the state of being poor, we would expect that all poor families with children would receive grants, while all non-poor families would be left out of the program, thus creating a short-cut for distinguishing eligible and non-eligible families. However, some evidence on the

[2] These values are valid for 2012. In 2006, the principal inclusion criteria was having an income per capita below R$ 120,00.

[3] The main implication here is that, since the data only shows us the beneficiary household and not the individual beneficiaries, we cannot differentiate between normal beneficiary teenagers in a household and those who are living in a beneficiary household but have decided to leave the program. Therefore, we will split the sample into younger and older children, with older children representing anyone between 10 and 15 years old.

implementation and institutional structure of “Bolsa Família” points to a probable mistargeting of grant allocation:

a) Institutional structure: even though legislation affirms that the benefit is not a vested right (Decree no. 5209, Art. 21, 17. September 2004), the law regarding the revision of household eligibility was only introduced in 2008, when a clause stated that revision of recipients should be done every two years (Decree no. 6392, Art.21, 2008). Furthermore, Art. 21 § 1 states that even if per capita income varies and infringes the eligibility criteria, there is no automatic termination of the program. As stated in Art. 21§ 1 (Decree no. 6392, 2008), exclusion happens only if it is proven that there is any omission or false information provided in the register which would falsely qualify the family for the program, the program beneficiary works in a paid elective position in any of the three government spheres, or if a family voluntary leaves the program.
b) Program development: In 2003, three cash transfer programs were merged into “Bolsa Família”. All recipients of the previous programs automatically became beneficiaries of “Bolsa Família”. This merging process was implemented without any eligibility checks, and thus generated problems clarifying the eligibility of recipients.
c) Implementation: Handa and Davis (2006) state that the self reported assignment variable “per capita income” is information that is not verified. Concerning the question’s formulation in the CADÚNICO, it seemed to be imprecise. Other programs present a proxy means test based on a set of variables which are less susceptible to manipulation (such as those in Ecuador and Jamaica). An additional feature is the household visits as important method for verification. In rural Mexico those visits were an integral element in program implementation, but not conducted in Brazil. (Handa and Davis, 2006)
d) Reported income: A large share of income in poor households in Brazil comes from informal work. As income is self reported and poorly checked, and informal work cannot be tracked by the agency, combined these factors might well be a source of grant allocation bias.
e) Self selective program design
f) Decentralized targeting and monitoring system at municipality level.

Combined, these factors are a good indication that grants might be misallocated in the country. As stated by Handa and Davis (2006), "of all conditional cash transfer programs Brazil's Bolsa Família and its predecessors seem to be the most susceptible to beneficiary manipulation and measurement error.“ [4]

So, although we understand the main structure of the program, several factors complicate the analysis. Firstly, the eligibility criteria are self-reported and non-checked information; secondly, exclusion from the program does not necessarily occur when there are changes in per capita income, meaning that the eligibility threshold can be crossed; finally, the merging process from 2004 made each additional recipient automatically eligible for the BFP, without sufficient checking of information. Based on this, the program's targeting can be considered very deficient, and is possibly the main cause of a contamination problem among the eligibility cut-off range. This will be the most important problem we will face when discussing the framework. [5]

## 2 Concepts and method

This session is ment to provide a quick discussion about potential methodology for the analysis of the "Bolsa Família" Program, taking into account the data we dispose and the particular problems we face around program implementation.

When analyzing policy measures with non-experimental studies, we encounter the evaluation problem. Ideally, we would look at how a non-treated individual would perform if the same individual would have received treatment (Heckman et al. 1999:1877pp.). However, there is no possibility of observing this effect for the same person, i.e., once with treatment and once without. In a treatment evaluation setting, this problem is solved by searching for a comparable counterfactual group, whose only difference is the participation in the program. Hence, socio-economic characteristics need to be comparable between the treated and non-treated person.

[4] We do recognize that monitoring is passing through significant changes in latest years. In December 2006 a new monitoring system through satellites and internet was inaugurated permitting the information flow between local communities and the government. The Ministry of Education is also supposedly following up educational conditionalities. However we can't say to which extend these improvements are affecting non-compliance of conditionalities and beneficiary manipulation.

[5] It is worth noting, however, that Castañeda et al. (2005) find assignment methods do in fact efficiently select extremely poor people, but they are is not efficient in excluding non-eligible individuals.

Given a program design where the inclusion criterion is conditional on per capita income, one possible methodological approach would be Regression Discontinuity Design (RDD). RDD is a quasi-experimental framework, which is applied for program evaluation in the case of non-experimental data. The major feature that characterizes RDD is that the data illustrates a discontinuous function of a particular assignment variable (or set of variables), as well as creating a threshold differentiating eligible and non-eligible individuals. Commonly, the performance of this discontinuous function is determined by a program regulation. RDD therefore uses this pattern of discontinuous dependence created generally by administrative regularization, and applies the eligibility criteria as an instrumental variable to estimate the average treatment effect (ATE).

The discontinuity pattern differs based on the participation function and presents two different types: the "sharp", and the "fuzzy" design. A deterministic participation function of *z* presents a sharp design, while a non-deterministic function of *z* leads to a fuzzy design. Both present the discontinuous cut-off point at the assignment variable. RDD has the ability of taking observed and unobserved heterogeneity into account for a given range around the assignment threshold. This is based on the argument that around the cut-off point, treated and non-treated individuals are similar (no observed or unobserved heterogeneity).

As discussed earlier, we should expect a sharp discontinuity around the eligibility cut-off point (R$120)[6], based on program design; however, program implementation lacks meticulous execution. Even though we considered using RDD, contamination between both groups refrained us from using it (See Figure 1 and 2)[7].

Alternatively, Propensity Score Matching (PSM) is a framework which aims to replicate an experimental data set by analyzing this comparable group. It matches two individuals, one treated and one non-treated, with both having the same probability of receiving treatment, conditional on the vector of characteristics X (Rosenbaum and Rubin 1983). Conditional on the following assumptions, PSM finds that the non-treated individual fulfills the same characteristics as the treated one, and compares the outcome between the two groups.

[6] Assignment level in 2006.

[7] However, we implemented a two stage IV regression following the "fuzzy" RDD design. The regressions presented the same effect as the PSM, but without significant effects.

1. Conditional Independence Assumption (CIA) or balancing property:

   $y_i^0 \perp d_i \,|P\,(X_i)$, where $y_i^0$ is the potential outcome of individual i in the non-treated group, $d_i$ stands for the treatment of individual i, and P($X_i$) is the probability of getting treatment conditional on the vector of observable characteristics X. The assumption certifies that the vector of explained variables X fully determines the possible outcome within the non-treated group, $y^0$, hence being able to overcome treatment selection bias. In short, it means that based on Xs, we have all relevant variables which may be able to determine the participation rule regarding treatment, as well as the non-treated outcome. Thus, after controlling for Xs, we can say that the outcome of individuals without treatment ($y_i^0$) is independent of their participation status.[8]

The balancing process used for the propensity score makes the treated and non-treated individuals comparable to each other, assuring that the CIA holds.

2. Overlap (or common support) condition assumption: 0<Pr [d=1|x] < 1. This assumption certifies that for each characteristic x, there are individuals from the treated and non-treated group, indicating an overlapping trend between both.

Once the propensity score is calculated, matching is performed based on fitting weights assigned to the neighborhood of P(X). Conditional on the assumptions above, the matching estimator is then:

$\hat{\alpha}^M = \sum_{i \in T} \{ y_i - \sum_{j \in C} \widetilde{w_{ij}} y_i \} \; w_i$, where T and C stand for the treatment and control group, i for the treated individuals in the treatment group T, j for non-treated individuals in the control group C, $\widetilde{w_{ij}}$ for the weights on j for group i, while $w_i$ stands for the reweighting that is needed to build the distribution for the beneficiary group.

PSM has the ability to address program impact in a broader way (even though observations might be dropped in the cause of common support). It is also capable of increasing the likelihood of comparisons across the treated and counterfactual groups, and could lower the bias of a program's impact in the case of a large common support (Khandker et. al, 2010). However, if unobservable characteristics do, in any way, determine program participation, the conditional independence assumption[9] will not hold, leading to a considerable program selection bias. RDD is able to overcome

[9] This assumption needs to hold for an analysis of the average treatment effect on the treated (ATT). In the case of the average treatment estimator (ATE), this would also need to hold for the treated sample. However, we focus on ATT.

unobservable heterogeneity given a particular range for an assignment variable, while PSM provides us with complete sample validity of a program's effect, but is not able to get rid of unobservable selection factors.

Cameron and Trivedi (2005:872) listed the cases in which this matching technique is a good methodology. First, it is important to have a good set of variables Xs to build a good control group. Blundell and Costa Dias (2008) argue that this choice is a delicate issue, as by choosing a large set of X variables, the overlapping support condition will not be fulfilled, and a small set would lead to a problem of unbalance where the CIA doesn't hold. Additionally, unsuitable characteristics for X would create problems in both assumptions. Therefore, the authors suggest that the X vector should explain both the treatment upon its assignment and its outcome. This is the condition that leads to a balanced CIA assumption, but it is not necessarily the condition that allows the overlapping support condition to hold. **Second,** it is important to have a large amount of observations which we can control for. **Third,** the Average Treatment on the Treated (ATT) is the parameter of interest as "if selection bias from unobserved characteristics is likely to be negligible, then PSM may provide a good comparison with randomized estimates" (Khandker et. al, 2010).

Diaz and Handa (2006) provide a valuable contribution to the discussion about the use of PSM for the evaluation of CCT – programs. They test the reliability of PSM as an alternative framework for evaluating programs that were not implemented as randomized control trial. They compare outcomes from a non-experimental dataset (the Mexican national household survey on income and expenditure, ENIGH) to outcomes from experimental datasets, finding no statistically significant bias between the estimates, which indicates that the PSM is a good alternative tool when working with non-experimental data to carry out treatment evaluation. Additionally, they also argue that PSM is a good method if the researcher possesses a dataset containing a rich set of variables able to estimate the propensity score, as well as in-depth information on the beneficiary selection process. All in all, these findings strongly point towards the use of the evaluation of the "Bolsa Família" Program.

### 2.1.1 Matching Algorithm

We base our findings on four matching algorithms: One-to-one matching (OO), nearest neighbor (NN), caliper and kernel.[10]

One-to-one matching consigns a weight of 1 if the score of the treated individual is the most nearby standing neighbor to the non-treated individual, and zero otherwise, so that each treated observation is matched with one comparable non-treated observation. Nearest neighbor matching is a less strict application and uses a certain range for non-treated individuals which are the K- closest match. This method has one drawback in that it could introduce bad comparisons to the matching process.

Radius matching was proposed by Dehejia and Wahba (2002), and is a modification of caliper matching. Caliper matching chooses the counterfactual group in one particular range of the propensity score only, providing an alternative to overcoming bad matches (Blundell and Costa Dias 2008). The radius consists of using the set of individuals in the control group within the caliper.

Kernel matching is a non-parametric matching estimator. The method calculates weights according to the distance in the propensity scores between treated and non-treated individuals, where the weight assigned to the treated individual tends closer to one as the kernel function draws nearer to the matched individual, while it falls as the propensity score of the matched observation becomes farer. Hence, kernel matching is a weighted regression of the counterfactual group outcome on an intercept, with weights given by the kernel weight (Smith and Todd 2005).

### 2.2 Dataset

The PNAD (Pesquisa Nacional de Amostra de Domicilios – the national household survey) was created in 1967, becoming an annual publication in 1971. Since then, there has been a coverage amplification process. In 2004, the last areas lacking coverage were covered; rural areas from Rondônia, Acre, Amazonas, Roraima, Pará and Amapá (northern districts). The questionnaire presents sections regarding population mobility (migration), education, fertility, health, the labor market, child labor, social insurance coverage, income, living conditions and asset possession. The survey also has supplementary sections which vary each year; in 2004 and 2006 questions regarding cash transfers to households were included, for example. Thus, we will mainly

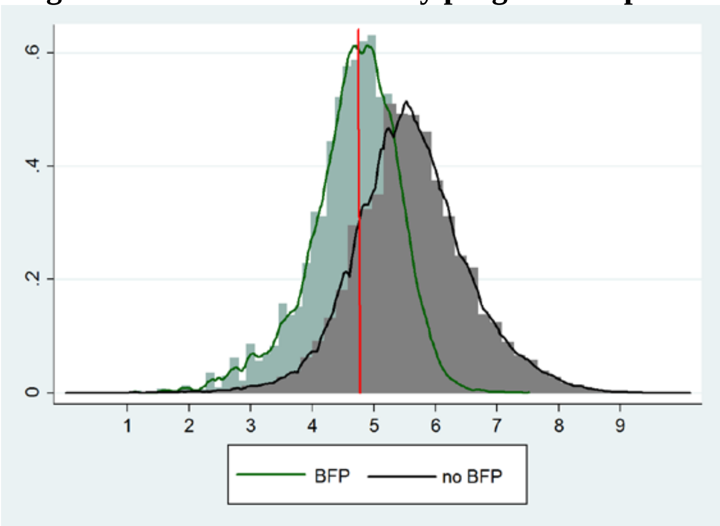
[10] Results for kernel analyzes were only reported for the sub-sample analysis. Local linear analyzes were also used but are not reported, since estimators rarely present big differences between matching methods overall.

concentrate on the PNAD 2006, only using the PNAD 2004 to analyze how the program evolved over these two years. The PNAD 2006 collected data from 410 241 individuals and 145 547 households and the survey covered all 27 districts, with a cross section survey being conducted every year in September. For our analysis, we only use only school children aged between 6 and 17.

## 2.3 Targeting

Here, we quickly go through some descriptive statistics, with the main intention being to analyze the program's targeting scheme. Figure I show the distribution of PCI for beneficiary and non-beneficiary families with children. Here, we can see an overlapping trend around the eligibility threshold instead of a clear cut-off. Even though this pattern goes against the use of RDD and clearly points towards mistargeting, it is important for the success of the PSM.

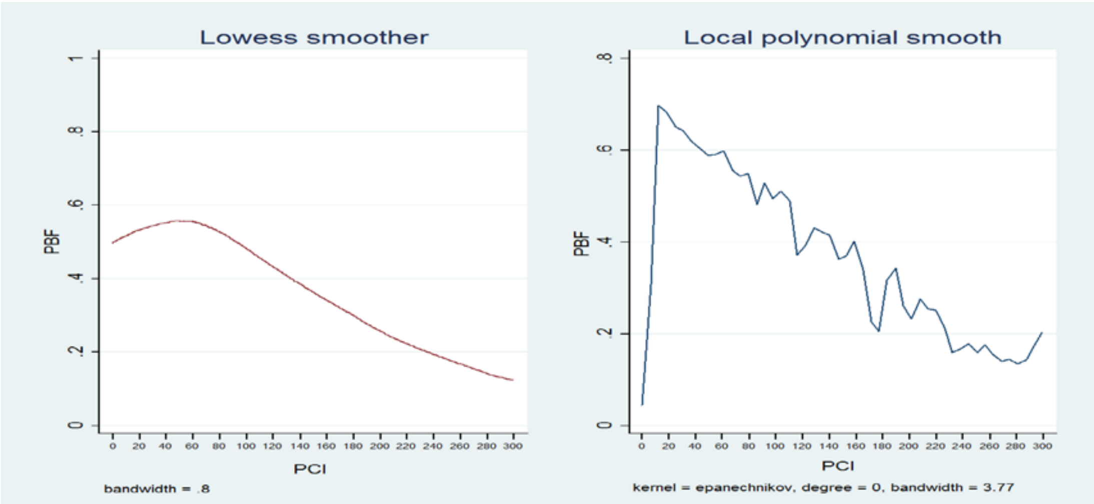
**Figure 1 PCI distribution by program recipient**




Source: PNAD 2006, own calculation

Figure 2 looks at the probabilistic relationship between being a recipient and PCI. On the left, we can see the locally weighted regression, and on the right a local polynomial regression. In both graphs, we can see that even though the probability of being a recipient falls as income increases, there is no clear downward fall around the eligibility threshold.

**Figure 2 Relationship between Treatment and PCI**




Source: PNAD 2006, own calculation

We will now focus on the application of PSM.

# 3 Application and Findings

### 3.1.1 Calculating the Propensity Score

The first step of propensity score matching is essential for the overall methodology. In order to calculate the probability of receiving treatment, we implemented a probit model as for the binary treatment case, the application of a logit or a probit model achieves analogous results (Caliendo and Kopeinig 2008).

Of greater concern is the choice of variables for the calculation of the propensity score. Since the CIA is a prerequisite for PSM, the propensity score should be composed in a way that the CIA holds, meaning that the outcome variable has to be "independent of treatment conditional on the propensity score" (Caliendo and Kopeinig 2008). Here, we then deal with two particular options; one is to introduce as many variables as possible with the argument that overparameterizing does not affect consistency or cause bias in the estimates; the second is to introduce only relevant variables for the determination of treatment.

Bryson, Dorsett, and Purdon (2002) present two arguments against the introduction of excessive parameters. Overparameterizing could compromise the common support condition by downsizing the common support area. Secondly, we may encounter an increase in variance.

Rubin and Thomas (1996) state that for prudence factors, the introduction of extra variables is justified. When doubts regarding the relevance of the covariate, or its relation with the outcome exist, the authors clearly recommend that all possible relevant variables should be introduced in the estimation. Caliendo and Kopeinig (2008) argue that if this approach is to be used, the selection process regarding the variable should rely on economic theory over empirical findings and the institutional setting.

However, what is important to emphasize is that the propensity score is not designed to estimate selection into treatment, but instead it is a simplifying tool for the multidimensionality problem. Hence, we followed the recommendation of Rubin and Thomas (1996) and introduced variables which, despite having a questionable level of significance and respective importance, were introduced into the propensity score model due to precautionary reasons and having a better fit. For the propensity score, we follow the specification of Dehejia and Wahba (1999), which requires the interaction terms and higher order terms.

To calculate the propensity score we introduced the following variables as independent variables for program participation[11]:

- Income variables: asset index.[12]
- Characteristics of the child: race (four dummies for being indigenous, black, Asian or mulatto), age, birth registered.
- Characteristics of the household head: head age and head age squared, the interaction of head age and eligibility[13], head is literate, head is female,

[11] The model was calculated with the psmatch2 command following Leuven and Sianesi (2003).

[12] The Asset Index was calculated through a polychoric PCA (Olsson, 1979). The polychoric and polyserial correlation is the estimate of the maximum likelihood of the correlation between the continuous asset variables in the index. This methodology first calculates the polychoric correlation and then runs the traditional PCA. Once the maximum likelihood is calculated for the correlation matrix, it is consistent, asymptotically normal and asymptotically efficient. (Kolenikov and Angeles, 2004). Included variables you find in Appendix 2.

[13] Ai and Norton (2003) show that the introduction of interaction terms in logit or probit models cannot be analyzed merely by looking at the sign of the estimator, its magnitude and statistic significance. The authors propose an alternative method for this case and show, with an example, that if the interaction term is not calculated correctly, the results may present an incorrect inference. Since I am not interested in predicting program participation, I decided to introduce the interaction term as it is, acknowledging the fact that it may present an incorrect inference.

education years of the household head (linear and squared), and type of work executed by the head.[14]

- Characteristics of the mother: mother alive, mother lives in the household.
- Characteristics of the household: 26 state dummies[15], 8 rural/urban dummies, 2 area dummies ("favela" and "aldeia")[16], number of children, number of members in household.
- Intertemporal variable: age of the oldest child

#### 3.1.1.1 The Common Support Condition

The overlapping pattern of the data is a particular and necessary characteristic of the PSM method, and is determined based on the propensity score.

**Table 1 Region of Common Support**

| | off Support | on Support | Total |
|---|---|---|---|
| Complete Sample (6 - 17 years old) | | | |
| Untreated | 1,657 | 55,903 | 57,560 |
| Treated | 27 | 29,951 | 29,978 |
| Total | 1,684 | **85,854** | 87,538 |
| Teenager Sample (10-15 years old) | | | |
| Untreated | 1,704 | 26,815 | 28,519 |
| Treated | 14 | 16,024 | 16,038 |
| Total | 1,718 | **42,839** | 44,557 |
| Children (6-9 years old) | | | |
| Untreated | 979 | 17,885 | 18,864 |
| Treated | 29 | 9,960 | 9,989 |
| Total | 1,008 | **27,845** | 28,853 |

Source: PNAD 2006, own calculations

Table 1 shows the area of common support, conditional on the 71 variables introduced for the calculation of the selection into treatment, i.e., propensity score. This ensures that there are no observations among those treated whose propensity score is higher than the maximum, or less than the minimum, of the counterfactual group. Since the matching will always be performed on the support area, we will work with 85 845 observations, with 55 905 untreated and 29 951 treated children aged between 6 and 17.

[14] In the agriculture, transformation, construction, commerce, food, transport, public service, social service, and domestic industries, as well as other industry, services, work and undefined groups.

[15] Even though recognizing its importance, we weren't able to introduce municipality dummies in the PSM, since the IBGE (brasilian institute of geography and statistics) strictly prohibit the access to the id code at this level. However, we do consider that, if municipality is uncorrelated with the outcome and if these unobservables are random distributed over the states, then this bias should also be random distributed, not affecting our estimates.

[16] Favelas are the regional expression for slums, and aldeia is an expression for indigenous villages.

The common support condition holds for all other sub samples. Regarding the concern expressed by Bryson, Dorsett, and Purdon (2002) overparameterizing did barely constraint the common support area.

#### 3.1.1.2 The Balancing Condition

The balancing condition is the prerequisite to using the propensity score as the matching technique by showing that treated and non-treated are similar and comparable.

The output of the balancing shows the t-test for the equality of the means in both the treated and non-treated groups for the matched and the unmatched samples. The test relies on the regression of a particular variable on the treatment indicator, and is done in the support area. The results show the standardized bias, as well as the reduction of this bias as a percentage due to matching.

The output of the balancing test for the whole sample can be found in appendix 4 for the OO matching. Here, it is important to verify the balancing condition for each variable after the matching process, in order to see if variables' means of non-treated are comparable to means of treated ones. We can see that after matching, most of the treated and non-treated groups become comparable by the different variables. Furthermore, the percentage of bias reduction is high for most of them, indicating that the balancing condition holds.[17]However, some variables are not balanced. Instead, they even present an increase in estimation bias. For a robustness check, we run the estimations without variables that can over- or underestimate our results, such as the ones presenting high percentage of bias increase (indian, female head, some area and state dummies, for example) and even the asset index, since after matching, the control group is significantly poorer than treated. We find no substantial differences in our estimates, however. The effect of the treatment varies in ± 0.5% in response to these robustness checks, while the exclusion of the asset index slightly decrease treatment effect on 1%. We hence decide to include all variables considered important, mainly because they represent a low percentage of the bias and are considered to be important for the fit.

[17]We do find however, that for some variables, treated and non-treated are still statistically significantly different after matching. Even though this can be caused by a possible unobservable bias, we consider this to be a result of the high amount of observations in the dataset, making any difference statistically significant.

## 3.2 Findings

We will now analyze the results of the PSM on enrollment and attendance. For robustness, we implemented four matching methods: one-to-one matching, nearest neighbor, radius matching and kernel matching. The specification of the matching models is as follows: for one-to-one matching, the counterfactual group was constructed with replacement; for nearest neighbor matching, we use five nearest neighbors; we used a caliper of 0.01 for the radius matching, and for kernel matching and local linear matching, the bandwidth is 0.06. In the following, the results between 2004 and 2006 are compared.

### 3.2.1 Comparison between 2004 and 2006

We looked at the program's effect on two different variables, enrollment and attendance, at two different points of time, 2004 and 2006.

Table 2 shows the enrollment rates by year and treatment, as well as the average treatment effect on the treated (ATT) on enrollment for the complete sample of children aged between 6 and 17.

**Table 2 Effects on Enrollment: 2004 and 2006**

| Bolsa Família effect on enrollment | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Matching | 2004 | | | | | 2006 | | | | | 2004/2006 | | |
| | 92.96 | | | | | 93.12 | | | | | -0.16 | | |
| | treated | controls | diff | s.e. | t-stat | treated | controls | diff | s.e. | t-stat | diff | | t-stat |
| unmatched | 94.33% | 93.85% | 0.48% | 0.00 | 2.33 | 93.15% | 93.46% | -0.31% | 0.00 | -1.74 | 1.18% | | |
| One to one | 94.33% | 88.73% | 5.60% | 0.00 | 14.28 | 93.15% | 88.47% | 4.68% | 0.00 | 13.04 | 0.92% | ** | 1.73 |
| NN | 94.33% | 89.26% | 5.07% | 0.00 | 17.54 | 93.15% | 88.65% | 4.51% | 0.00 | 16.40 | 0.57% | * | 1.42 |
| radius | 94.33% | 89.28% | 5.06% | 0.00 | 20.39 | 93.15% | 88.50% | 4.65% | 0.00 | 19.40 | 0.41% | | 1.18 |

Source: PNAD 2004 and 2006, own calculation

We can see that 92.96% of children were enrolled in 2004. This number slightly increased in 2006, with a total enrollment rate of 93.12%. In the case of cash transfer recipients, the enrollment rate was slightly higher at 94.33%, compared to that of non-recipients at 93.85%. In 2006, however, enrollment rates for treated children decreased by 1.18% compared to 2004.

Regarding the program's impact on enrollment, we can see that the ATT was around 5% higher for treated children in 2004 using the different matching estimators, while in 2006, even though it was positive and significant, it slightly decreased to around 4.5%.. Two estimators showed this decrease one-to-one matching (t-stat: 1.73) and NN matching (t-stat: 1.42) present significant results for the differences between periods. This result is mainly driven by the already reported fall in the probability of recipient enrollment in 2006.

**Table 3 Effects on Attendance (Schooldays missed): 2004 and 2006**

| Bolsa Família effect on attendance | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Matching | 2004 | | | | | 2006 | | | | | 2004/2006 | |
| | 2.40 | | | | | 1.76 | | | | | 0.64 | |
| | treated | controls | diff | s.e. | t-stat | treated | Controls | diff | s.e. | t-stat | diff | t-stat |
| unmatched | 2.51 | 2.26 | 0.25 | 0.05 | 4.73 | 1.86 | 1.70 | 0.16 | 0.03 | 5.79 | | |
| One to one | 2.51 | 2.85 | -0.34 | 0.10 | -3.58 | 1.86 | 2.15 | -0.28 | 0.06 | -5.05 | -0.06 | -0.51 |
| NN | 2.51 | 2.87 | -0.36 | 0.07 | -4.83 | 1.86 | 2.21 | -0.35 | 0.04 | -8.03 | -0.01 | -0.51 |
| radius | 2.51 | 2.86 | -0.36 | 0.07 | -5.44 | 1.86 | 2.19 | -0.32 | 0.04 | -8.27 | -0.04 | -0.46 |

Source: PNAD2004 and 2006, own calculation

Overall, we can see that the program had a positive impact on enrollment, with a 4-5% higher probability of being enrolled for program recipients in both periods. Nevertheless, our results also show that enrollment rates for treated children fell over time by 1.18%, even though enrollment rates slightly improved between 2004 and 2006. This evidence was also found when analyzing treatment effects over different groups. Appendix I offers a quick overview of the results for heterogeneous effects in the year 2004 (which will be carefully analyzed in further detail). Here, we can see that, besides the increase in enrollment rates between younger treated children (95% in 2004 and 96% in 2006), the estimated fall from 2004 to 2006 is one percentage point. This effect seems to be more or less constant over groups. When looking at regional effects, besides the Midwest region, the enrollment rate fell by one percentage point in 2006. When genders are compared, the enrollment rate for treated males was 93% in 2004, falling to 92% in 2006, while for females, enrollment decreased from 95.6% to 94% during this period. This occurred in both rural and urban areas, where it fell from 93.5% to 92%, and 94.5 % to 93.5%, respectively.  This result might therefore indicate that, over time, the effect of cash transfers might fade.

Table 3 shows the results for school attendance in 2004 and 2006. While children missed on average of 2.4 days of school over the preceding two months in 2004, this number decreased by 0.64 days in 2006, i.e., 1.76 days missed overall, on average.

In the case of recipient children, they missed slightly more school then the average (around 0.1 days) for both years. However, when compared to their counterfactual group, cash transfers are estimated to have a positive impact on attendance of 0.35

fewer days missed for 2004, and around 0.32 fewer for 2006. We didn't find any significant differences in attendance over time.

## 3.3 Heterogeneous Treatment Effects

While finding an overall effect of "Bolsa Família", we are also interested in the differences in its impact differences between age groups, regions, genders and areas (rural/urban). The next section will concentrate on these heterogeneous treatment effects for the year 2006.[18]

### 3.3.1 Heterogeneous treatment effects by age

Table 4 shows the results of cash transfers for two different age groups. First of all, we noticed that enrollment rates were different when age groups were compared. While younger children, when treated, had enrollment rates of around 96%, and 92% when they were not treated, older children showed, on average, a four percentage point lower probability of being enrolled. When analyzing differences in the program's impact by age, the effect of cash transfers does not seem to differ between groups; both groups, older and younger, show a 4% higher probability of being enrolled when treated. The results vary slightly here when matching methods were compared. Our findings point to the fact that "Bolsa Família" is not able to eliminate the fact that older children are more prone to drop out of school as they get increasingly older.

We can identify some differences in the effect of the program on attendance, however. On average, the ATT estimator for younger children varies between -0.38 and -0.54, depending on the matching estimators. This can be interpreted as a positive impact on attendance due to program participation. For older children, the effect of program participation was also positive on attendance. Teenagers missed around 0.28 less days of school than the counterfactual group. Therefore, I find a higher impact on attendance among younger children; younger (treated) children miss, on average, around 0.2 less days of school when compared to older (treated) children. This effect seemed to be mainly driven by the fact that (non-treated) younger children miss, on average, slightly more days of school than (non-treated) older children (around 2.30 days for younger, and 2.14 days for older children), and that cash transfers are able to pull this effect up to the same level for both age groups (around 1.86 missed days).

[18] For results on the common support area, please see appendices 7, 9 and 10

**Table 4 Heterogeneous treatment effects by age**

| Heterogeneous shocks by age group | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Matching | Enrollment | | | | | Attendance | | | | |
| | treated | controls | diff | s.e. | t-stat | treated | controls | diff | s.e. | t-stat |
| 6-9 years | | | | | | | | | | |
| One to one | 96.00% | 91.99% | 4.02% | 0.005 | 7.62 | 1.85 | 2.39 | -0.54 | 0.094 | -5.79 |
| NN | 96.00% | 92.17% | 3.83% | 0.004 | 9.99 | 1.85 | 2.30 | -0.45 | 0.071 | -6.30 |
| radius | 96.00% | 92.32% | 3.68% | 0.003 | 11.25 | 1.85 | 2.23 | -0.38 | 0.064 | -5.89 |
| Kernel | | | | | | | | | | |
| 10-17 years | | | | | | | | | | |
| One to one | 91.73% | 87.47% | 4.26% | 0.005 | 9.09 | 1.87 | 2.14 | -0.27 | 0.070 | -3.82 |
| NN | 91.73% | 87.58% | 4.15% | 0.004 | 11.30 | 1.87 | 2.14 | -0.27 | 0.055 | -4.94 |
| radius | 91.73% | 87.70% | 4.03% | 0.003 | 12.46 | 1.87 | 2.15 | -0.28 | 0.050 | -5.61 |
| Kernel | | | | | | | | | | |
| ttest for statistically significantyl differneces: Primary-Secondary (6-10 vs 10-17) | | | | | | | | | | |
| | diff | | t-stat | | | diff | | t-stat | | |
| One to one | 0.00 | | -0.35 | | | -0.27 | | | -2.33 | *** |
| NN | 0.00 | | -0.59 | | | -0.18 | | | -1.96 | ** |
| radius | 0.00 | | -0.76 | | | -0.10 | | | -1.20 | |

Source: PNAD 2006, own calculation

This results shows that the program has a positive impact on attendance for both age groups, while the size of the impact seems to be higher for younger children then for older ones.

### 3.3.2 Heterogeneous treatment effects by region

The level of poverty and inequality in Brazil is known to be very heterogeneous across regions. While the HDI of the whole country was 0.803 in 2006, the HDIs of the North and North-East were 0.733 and 0.772, respectively. The South, South-East and Midwest presented significantly higher development indicators, at 0.837, 0.835 and 0.824, respectively (BCB 2009). Hence, we are interested if the program's impact has different outcomes across different regions, and thus whether it is able to overcome regional inequalities.

Table 5 shows the effect of the program in the five regions of Brazil: North (N), North-East (NE), South (S), South-East (SE) and Midwest (MW), focusing on the effect of "Bolsa Família" in the North and North-East regions.

Overall, we find a significant positive effect of cash transfers on enrollment in all regions. In the North, recipients had a 4% higher probability of being enrolled then non-beneficiaries. Additionally, this number was 5% in the North-East, and around 2% in the South-East, South and Midwest. We also found positive effects on attendance for all regions besides the Midwest, varying from 0.2 to 0.3 fewer missed days of school. When focusing on the northern region, we can see that enrollment rates are around 92%

**Table 5 Heterogeneous shocks by region**

| Heterogeneous shocks by region | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Matching | Enrollment | | | | | Attendance | | | | |
| | treated | controls | diff | s.e. | t-stat | treated | controls | diff | s.e. | t-stat |
| North | | | | | | | | | | |
| | 91.45 | | | | | 1.65 | | | | |
| no Matching | 92.21% | 91.40% | 0.81% | 0.00 | 1.64 | 1.87 | 1.54 | 0.33 | 0.08 | 4.33 |
| One to one | 92.19% | 88.49% | 3.70% | 0.008 | 4.49 | 1.87 | 2.13 | -0.26 | 0.128 | -2.02 |
| NN | 92.19% | 88.43% | 3.77% | 0.006 | 5.90 | 1.87 | 2.09 | -0.22 | 0.102 | -2.12 |
| radius | 92.19% | 88.00% | 4.19% | 0.006 | 7.27 | 1.87 | 2.08 | -0.21 | 0.092 | -2.26 |
| Kernel | 92.19% | 88.02% | 4.17% | 0.006 | 7.30 | 1.87 | 2.11 | -0.24 | 0.092 | -2.56 |
| North-East | | | | | | | | | | |
| | 92.76 | | | | | 1.88 | | | | |
| no Matching | 93.72% | 92.42% | 1.30% | 0.00 | 4.40 | 1.88 | 1.88 | 0.00 | 0.05 | -0.02 |
| One to one | 93.72% | 89.40% | 4.32% | 0.006 | 7.52 | 1.88 | 2.09 | -0.21 | 0.096 | -2.23 |
| NN | 93.72% | 88.69% | 5.03% | 0.004 | 11.18 | 1.88 | 2.23 | -0.21 | 0.078 | -4.50 |
| radius | 93.72% | 88.45% | 5.27% | 0.004 | 12.91 | 1.88 | 2.22 | -0.21 | 0.072 | -4.68 |
| Kernel | 93.72% | 88.49% | 5.23% | 0.004 | 13.07 | 1.88 | 2.19 | -0.21 | 0.071 | -4.40 |
| South-East | | | | | | | | | | |
| | 94.45 | | | | | 1.84 | | | | |
| no Matching | 93.00% | 95.06% | -2.06% | 0.00 | -5.85 | 1.87 | 1.82 | 0.05 | 0.06 | 0.92 |
| One to one | 93.10% | 91.12% | 1.98% | 0.007 | 2.73 | 1.86 | 2.18 | -0.32 | 0.108 | -2.94 |
| NN | 93.10% | 90.71% | 2.39% | 0.006 | 4.27 | 1.86 | 2.16 | -0.29 | 0.082 | -3.53 |
| radius | 93.10% | 90.70% | 2.40% | 0.005 | 4.99 | 1.86 | 2.15 | -0.29 | 0.074 | -3.84 |
| Kernel | 93.10% | 90.71% | 2.39% | 0.005 | 5.03 | 1.86 | 2.15 | -0.28 | 0.073 | -3.83 |
| South | | | | | | | | | | |
| | 92.89 | | | | | 1.65 | | | | |
| no Matching | 90.86% | 93.65% | -2.79% | 0.01 | -4.79 | 1.85 | 1.60 | 0.25 | 0.08 | 3.13 |
| One to one | 91.03% | 89.83% | 1.20% | 0.008 | 1.07 | 1.85 | 1.98 | -0.14 | 0.155 | -0.88 |
| NN | 91.03% | 88.94% | 2.09% | 0.009 | 2.30 | 1.85 | 2.05 | -0.21 | 0.123 | -1.68 |
| radius | 91.03% | 88.97% | 2.06% | 0.008 | 2.57 | 1.85 | 1.97 | -0.14 | 0.112 | -1.20 |
| Kernel | 91.03% | 89.00% | 2.02% | 0.008 | 2.60 | 1.85 | 1.95 | -0.11 | 0.109 | -0.97 |
| Center-West | | | | | | | | | | |
| | 93.77 | | | | | 1.48 | | | | |
| no Matching | 94.25% | 93.98% | 0.27% | 0.01 | 0.45 | 1.73 | 1.41 | 0.33 | 0.09 | 3.79 |
| One to one | 94.31% | 93.26% | 1.05% | 0.009 | 1.16 | 1.75 | 1.67 | 0.08 | 0.144 | 0.56 |
| NN | 94.31% | 91.73% | 2.58% | 0.007 | 3.52 | 1.75 | 1.62 | 0.13 | 0.114 | 1.16 |
| radius | 94.31% | 91.56% | 2.75% | 0.007 | 4.21 | 1.75 | 1.58 | 0.17 | 0.107 | 1.59 |
| Kernel | 94.31% | 91.70% | 2.61% | 0.006 | 4.08 | 1.75 | 1.59 | 0.16 | 0.106 | 1.54 |
| ttest for statistically significantly differences | | | | | | | | | | |
| North-South | | | | | | | | | | |
| | diff | | t-stat | | | diff | | t-stat | | |
| One to one | 2.51% | | 2.15 | ** | | -0.12 | | 0.60 | | |
| NN | 1.67% | | 1.51 | * | | -0.01 | | 0.06 | | |
| radius | 2.13% | | 2.16 | ** | | -0.07 | | 0.51 | | |
| Kernel | 2.15% | | 2.22 | ** | | -0.13 | | 0.91 | | |

| | | | | |
|---|---|---|---|---|
| North-South East | | | | |
| One to one | 1.73% | 1.57 * | 0.06 | - 0.35 |
| NN | 1.38% | 1.62 * | 0.07 | - 0.56 |
| radius | 1.79% | 2.39 ** | 0.08 | - 0.64 |
| Kernel | 1.78% | 2.40 ** | 0.05 | - 0.39 |
| North-Center West | | | | |
| One to one | 1.05% | 2.16 ** | -0.34 | - 1.76 ** |
| NN | 2.58% | 1.22 | -0.35 | - 2.28 ** |
| radius | 2.75% | 1.66 ** | -0.38 | - 2.68 *** |
| Kernel | 2.61% | 1.82 ** | -0.40 | - 2.84 *** |
| North East -South | | | | |
| One to one | 1.20% | 3.11 *** | -0.08 | - 1.93 ** |
| NN | 2.09% | 2.89 *** | -0.01 | - 2.89 *** |
| radius | 2.06% | 3.57 *** | -0.08 | - 0.59 |
| Kernel | 2.02% | 3.66 *** | -0.11 | - 2.46 *** |
| North East-South East | | | | |
| One to one | 1.98% | 2.53 *** | 0.10 | - 3.68 *** |
| NN | 2.39% | 3.68 *** | 0.08 | - 4.46 *** |
| radius | 2.40% | 4.55 *** | 0.07 | - 4.82 *** |
| Kernel | 2.39% | 4.57 *** | 0.07 | - 4.85 *** |
| North East -Midwest | | | | |
| One to one | 1.05% | 3.04 *** | -0.30 | - 2.30 ** |
| NN | 2.58% | 2.84 *** | -0.35 | - 3.06 *** |
| radius | 2.75% | 3.28 *** | -0.39 | - 3.52 *** |
| Kernel | 2.61% | 3.47 *** | -0.38 | - 3.48 *** |

Source: PNAD 2006, own calculation

for treated children and 88% for the estimated counterfactual group. We find that the effect in the north is around two percentage points when compared to the S, SE and Midwest. This difference is statistically significant for the comparisons N-S, N-SE and N-Midwest. For attendance, we find some significant differences between the North and Midwest, with recipient children missing between 0.35 and 0.40 less days of school, on average, than recipient children in the Midwest. This effect is mainly driven by the fact that we didn’t find a positive effect of attendance in the Midwest.

In the case of the North-East, we can see that the heterogeneous shocks are even higher. When compared to the other three regions which present the highest disparities (S, SE, MW), the effect of cash transfers on enrollment is around two percentage points higher than in the other regions. This effect is significantly higher for the NE-S, NE-SE and NE–MW comparisons. Even though differences in attendance seem to be significant across the regions, they are also small and go in different directions in many cases - being higher when compared to the South and Midwest, and lower if compared to the South-East.

There are various reasons for the reported effects. While the North and North-East illustrate lower enrollment rates for non-treated children (88%), a disparity which is expected as this region has lower development indices, non-treated children in more developed regions present a higher probability of being enrolled ( 89% in the South, 91% in SE, and 92% in the Midwest). Thus, the heterogeneous effects we report here are mainly driven by a convergence in enrollment rates between non-treated children in low developed regions towards those of non-treated children in higher developed regions. Enrollment rates of “Bolsa Família” recipients in less developed regions are therefore closer to the enrollment rates of non-recipients in more developed regions. Consequently, the results here are evidence that cash transfers have been able to mitigate the gap between regions.

This effect is not so clear in the case of attendance, mainly because the variance across regions is rather small. However, what we find is that the effect on attendance in the Midwest is negative or inexistent. If treated, they miss more school than non-treated children.

### 3.3.3 Heterogeneous treatment effect by gender

Table 6 exhibits the differences between boys and girls. We can see that 93.42% of girls are enrolled, while boys display a slightly lower enrollment rate of 92.83%. When looking at treated boys and girls, we also noted that 94% of girls treated go to school, while only 92.3% of boys are enrolled. When looking at the program’s overall effect here, it seems to have had an equal impact on boys and girls; treated boys and girls both have a 4 % higher probability of going to school if compared to their counterfactual group.

**Table 6 Heterogeneous treatment effects by gender**

| Heterogeneous shocks by gender | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Matching | Enrollment | | | | | Attendance | | | | |
| | treated | controls | diff | s.e. | t-stat | treated | controls | diff | s.e. | t-stat |
| Male | | | | | | | | | | |
| | 92.83% | | | | | 1.82 | | | | |
| One to one | 92.31% | 87.51% | 4.80% | 0.01 | 9.18 | 1.94 | 2.30 | -0.37 | 0.08 | -4.64 |
| NN | 92.31% | 87.93% | 4.38% | 0.00 | 10.91 | 1.94 | 2.31 | -0.38 | 0.06 | -6.00 |
| radius | 92.31% | 87.94% | 4.37% | 0.00 | 12.57 | 1.94 | 2.28 | -0.34 | 0.06 | -6.18 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Kernel | 92.31% | 87.95% | 4.36% | 0.00 | 12.72 | 1.94 | 2.27 | -0.33 | 0.06 | -6.02 |
| Female | | | | | | | | | | |
| | 93.42% | | | | | 1.70 | | | | |
| One to one | 94.03% | 89.75% | 4.28% | 0.00 | 8.63 | 1.79 | 2.03 | -0.24 | 0.08 | -3.07 |
| NN | 94.03% | 89.92% | 4.11% | 0.00 | 10.81 | 1.79 | 2.03 | -0.24 | 0.06 | -3.96 |
| radius | 94.03% | 89.81% | 4.22% | 0.00 | 12.78 | 1.79 | 2.07 | -0.29 | 0.05 | -5.26 |
| Kernel | 94.03% | 89.81% | 4.22% | 0.00 | 12.89 | 1.79 | 2.06 | -0.27 | 0.05 | -5.00 |
| ttest for statistically signifcantly differences: Male-Female | | | | | | | | | | |
| | diff | | t-stat | | | diff | | t-stat | | |
| One to one | 0.52% | | 0.72 | | | -0.13 | | 1.14 | | |
| NN | 0.27% | | 0.48 | | | -0.13 | | 1.51 | | * |
| radius | 0.15% | | 0.30 | | | -0.06 | | 0.76 | | |
| Kernel | 0.14% | | 0.30 | | | -0.06 | | 0.82 | | |

Source: PNAD 2006 – own calculations

We also find that the effect on attendance was slightly higher, though not significantly different between girls and boys.

### 3.3.4 Heterogeneous treatment effects by area

Table 7 shows heterogeneous treatment effects according to rural, urban and slum areas. Here, we can see that 89.73% of children are enrolled at school in rural areas, while in urban areas, 93.91 % of children are enrolled. For treated children, the probability of being enrolled rises to 92.33% in rural areas, and is slightly smaller in urban areas (93.51%). We also found that the effect of “Bolsa Família” varies across regions - while it increases enrollment in rural areas by around 5.5 %, the positive effect in urban areas is estimated to be about 3.5 % for treated children (Table7). This difference of two percentage points between rural and urban areas is significant for all matching estimators. We also find significant differences regarding attendance. The effect of “Bolsa Família” on comparable recipients in rural areas is significantly higher than in urban areas. Rural recipients miss, on average, 0.5 less days than comparable recipients in urban areas (Table 7).

This result is mainly driven by the fact that non-treated children are anyhow more likely to go to school in urban then in rural areas. However, when treated enrollment probability from rural children become closer to enrollment probability of urban children. Hence, the program is able to overcome differences between regions and promote convergence in school enrollment.

The positive impact of cash transfers on enrollment in slum areas appeared to be 1% higher than in urban areas, though the difference is not significant. While we weren’t able to find significant effects on enrollment for urban and slum areas, we found that in the case of attendance, recipients living in slum areas miss, on average, slightly less school than comparable children living in other urban areas (0.06 days).

**Table 7 Heterogeneous treatment effect by area**

| Heterogeneous shocks by area | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Matching | Enrollment | | | | | Attendance | | | | |
| | treated | controls | diff | s.e. | t-stat | treated | controls | diff | s.e. | t-stat |
| Rural | | | | | | | | | | |
| | 89.73 | | | | | 1.88 | | | | |
| One to one | 92.33% | 86.42% | 5.91% | 0.01 | 5.71 | 1.81 | 2.54 | -0.73 | 0.17 | -4.24 |
| NN | 92.33% | 86.68% | 5.65% | 0.01 | 6.99 | 1.81 | 2.42 | -0.60 | 0.12 | -4.97 |
| radius | 92.33% | 86.82% | 5.51% | 0.01 | 7.14 | 1.81 | 2.41 | -0.60 | 0.11 | -5.24 |
| Kernel | 92.33% | 86.84% | 5.49% | 0.01 | 8.01 | 1.81 | 2.42 | -0.61 | 0.10 | -6.26 |
| Urban | | | | | | | | | | |
| | 93.91 | | | | | 1.73 | | | | |
| One to one | 93.51% | 90.23% | 3.28% | 0.00 | 8.71 | 1.88 | 1.98 | -0.09 | 0.06 | -1.52 |
| NN | 93.51% | 89.97% | 3.55% | 0.00 | 12.22 | 1.88 | 1.97 | -0.09 | 0.05 | -1.88 |
| radius | 93.51% | 90.02% | 3.50% | 0.00 | 13.93 | 1.88 | 1.98 | -0.10 | 0.04 | -2.25 |
| Kernel | 93.51% | 90.10% | 3.41% | 0.00 | 13.74 | 1.88 | 1.98 | -0.09 | 0.04 | -2.23 |
| Slums | | | | | | | | | | |
| | 91.63 | | | | | 1.95 | | | | |
| One to one | 92.99% | 88.29% | 4.70% | 0.02 | 3.07 | 2.00 | 2.27 | -0.27 | 0.23 | -1.16 |
| NN | 92.99% | 88.73% | 4.26% | 0.01 | 3.65 | 2.00 | 2.07 | -0.08 | 0.18 | -0.43 |
| radius | 92.99% | 88.73% | 4.50% | 0.01 | 4.27 | 2.00 | 2.13 | -0.15 | 0.17 | -0.86 |
| Kernel | 92.99% | 88.54% | 4.44% | 0.01 | 4.34 | 2.00 | 2.15 | -0.16 | 0.16 | -0.96 |
| ttest for statistically significantly differences | | | | | | | | | | |
| Rural-Urban | | | | | | | | | | |
| | | diff | | t-stat | | | diff | | t-stat | |
| One to one | | 2.63% | | 2.39 | *** | | -0.63 | | -3.48 | *** |
| NN | | 2.10% | | 2.45 | *** | | -0.51 | | -3.77 | *** |
| radius | | 2.01% | | 2.48 | *** | | -0.50 | | -4.13 | *** |
| Kernel | | 2.08% | | 2.85 | *** | | -0.52 | | -4.86 | *** |
| Slums-Urban | | | | | | | | | | |
| One to one | | 1.42% | | 0.66 | | | -0.18 | | -1.57 | * |
| NN | | 0.71% | | 0.98 | | | 0.01 | | -2.46 | *** |
| radius | | 1.01% | | 0.77 | | | -0.05 | | -2.21 | ** |
| Kernel | | 1.03% | | 0.85 | | | -0.06 | | -2.37 | *** |

Source: PNAD 2006, own calculation

# 4 Conclusion

CCT programs commenced in Latin America in the early 90's as an attempt to break the inter-generational cycle of poverty. Some major positive effects have been proven, and in this paper we particular focused on the effect of the Bolsa Família Program on enrollment and attendance. One would assume this should be automatic, but it is a condition that was at the point of analysis not efficiently monitored.

We find a positive impact of the BFP on education. Even though the program lacks strong monitoring, the probability of enrollment increases by four percentage points for children between the ages of 6 and 17, and estimates on attendance show that treated children miss 0.30 less days of school than untreated ones. The effect on enrollment slightly decreased from 2004 to 2006 by around 1%. Differences between regions were also found in 2006. The impact of the program is significantly higher in rural areas and in less developed regions, i.e., the north and north-east.

Although our findings point to positive effects on education, the estimated effect seems to be modest upon initial investigation and might not be considered sufficient enough to enhance educational quality and human capital. Furthermore, CCT programs might divert attention from necessary improvements to the ineffective public service, especially health, education and social insurance programs (Rawlings and Rubio 2005). Others raise concerns about the program's expansion as a political tool to ensure political support of the middle-class (Handa and Davis 2006). This has become one of the most common arguments in the country since speculations arose that President Luis Inácio Lula da Silva, elected in 2002, greatly expanded the program in 2004 with victory in the 2006 election in mind. If this is the case, controlling for the fulfillment of conditionality by recipients can be conducted in a low cost way, or even be removed altogether. This may be the case in Brazil, where controlling for recipient fulfillment of conditionality requirements is conducted by the municipalities "and is haphazard at best" (Handa and Davis 2006:10).

Even though "Bolsa Família" might be lacking in some areas, our paper provides evidence that it is indeed an engine of change. One of the main achievements of "Bolsa Família" its power to close the development gap between regions, with less developed and rural areas benefitting somewhat more from the grants. This can be seen through a convergence in enrollment rates between unequal regions. Furthermore, it is not the quantitative output estimated that matters, but instead the cash transfers, which we see as being able to promote structural changes in behavior, i.e., a change in commitment to education among the poor population.

## References


Ai, C. and Norton, E.C. (2003), “Interaction Terms in Logit and Probit Models”, *Economics Letters* 80, 123-129.

Barros, R. P. de; Foguel, M.N. and Ulyssea, G. (2006), *Desigualdade de Renda no Brasil - Uma Análise da Queda Recente*, Instituto de Pesquisa Econômica Aplicada, Brasilia, Brazil.

Behrman, J.R.; Sengupta, P: and Todd, P.E., (2005), “Progressing through PROGRESA - An Impact Assessment of a School Subsidy Experiment in Rural Mexico”, *Economic Development and Cultural Change* 54 (1), 237–75.

Boletim Regional do Banco Central do Brasil. Jan 2009, 91-94 Available in: http://www.bcb.gov.br/pec/boletimregional/port/2009/01/br200901b1p.pdf

Behrman, J.R.; Sengupta, P. and Todd, P.E. (2000), “The Impact of PROGRESA on Achievement Test Scores in the First Year”, Unpublished Manuscript, International Food Policy Research Institute, Washington DC.

Behrman, J.R.; Parker, S.W. and Todd, P.E. (2005), “Long-Term Impacts of the Oportunidades Conditional Cash Transfer Program on Rural Youth in Mexico”, Discussion Paper 122, Ibero-America Institute for Economic Research, Göttingen, Germany.

Blundell, R. and Costa Dias, M., (2008), “Alternative Approaches to Evaluation in Empirical Microeconomics”, Discussion Paper No. 3800, Institute for the Study of Labor (IZA), Bonn, Germany.

Bourguignon, F.; Ferreira F.H.G. and Leite, PG. (2003), “Conditional Cash Transfers, Schooling, and Child Labor - Micro-Simulating Brazil’s Bolsa Escola Program”, *World Bank Economic Review* 17 (2), 229-254.

Caliendo, M. and Kopeinig, S. (2008), “Some Practical Guidance for the Implementation of Propensity Score Matching”, *Journal of Economic Surveys* (2008) Vol. 22, No.1, pp. 31-72.

Cameron, A.C. and Trivedi, P.K. (2005), *Microeconometrics - Methods and Applications*, Cambridge University Press, New York.

Cardoso, E., and Portela Souza, A. (2004), “The Impact of Cash Transfers on Child Labor and School Attendance in Brazil.” Working Paper 04-W07, Department of Economics, Vanderbilt University, Nashville, TN.

Castañeda, T. and Lindert, K. (2005), “Designing and Implementing Household Targeting Systems - Lessons from Latin America and the United States”, Social Protection Discussion Paper 0526. World Bank, Washington, DC.

Decree no. 5209 and Decree no. 6392, 2008. Available in:<http://www.planalto.gov.br/ccivil_03/_ato2007-2010/2008/decreto/d6392.htm> and <http://www.planalto.gov.br/ccivil_03/_Ato2004-2006/2004/Decreto/D5209.htm#art24p>

Dehejia, R.H. and Wahba, S. (1999), “Causal Effects in Non Experimental Studies - Reevaluating the Evaluation of Training Programs”, *Journal of American Statistical Association* 94 (448), 1053-1062.

Dehejia, R.H. and Wahba, S. (2002), “Propensity Score Matching Methods for Non-Experimental Causal Studies”, *Review of Economics and Statistics* 84 (1), 151-161.

De Janvry, A.; Finan, F. and Sadoulet, E. (2006). “Evaluating Brazil’s Bolsa Escola Program: Impact on Schooling and Municipal Roles”. June 2006

de Janvry, A.; Finan, F; Sadoulet, E.; Vakis, R. (2006). “Can conditional cash transfer programs serve as safety nets in keeping children at school and from working when exposed to shocks?” Journal of Development Economics 79 (2006) 349 – 373. jan

Diaz, J.J. and Handa, S. (2006), “An Assessment of Propensity Score Matching as a Non-Experimental Impact Estimator Evidence from Mexico’s PROGRESA Program”, *The Journal of Human Resources* 41 (2), 319-345.

Bryson, A.; Dorsett, R. and Purdon, S. (2002), “The Use of Propensity Score Matching in the Evaluation of Labor Market Policies", Working Paper No. 4, Department for Work and Pensions.

Glewwe, P. and Olinto, P. (2004), “Evaluating the Impact of Conditional Cash Transfers on Schooling - An Experimental Analysis of Honduras”, Unpublished Manuscript, University of Minnesota, Minneapolis.

Glewwe, P. and Kassouf, A. L. (2008), “The Impact of the *Bolsa Escola/Família* Conditional Cash Transfer Program on Enrollment, Grade Promotion and Dropout Rates in Brazil”, ANPEC - Associação Nacional dos Centros de Pósgraduação em Economia (Brazilian Association of Graduate Programs in Economics).

Heckman J. J.; Lalonde, R. J. and Smith, J. A. (1999). *Handbook of Labor Economics*, Volume , Elsevier Science, Burlington.

IBGE (2004, 2006), Pesquisa Nacional de Amostra de Domicilios (PNAD 2004, PNAD 2006), IBGE, Rio do Janeiro.

Khandker S. R. Gayatri B. K., and H. A. Samad (2010). “Handbook on Impact Evaluation Quantitative Methods and Practices”. World Bank

Kolenikov, S. and Angeles, G. (2004), “The Use of Discrete Data in Principal Component Analysis - Theory, Simulations, and Applications to Socioeconomic Indices”, Working Paper of MEASURE/Evaluation Project No. WP-04-85, Carolina Population Center, University of North Carolina.

Leuven, E., Sianesi, B., 2003. PSMATCH2: stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate impalance testing. Available via Internet: < http://ideas.repec.org/c/boc/bocode/s432001.html>.

Macours, K. and Vakis, R. (2008), “Changing Households’ Investments and Aspirations through Social Interactions - Evidence from a Randomized Transfer Program in a Low-Income Country”, Unpublished Manuscript, Johns Hopkins University, Baltimore MD, and World Bank, Washington, DC.

Maluccio, J.A. and Flores, R. (2005) “Impact Evaluation of a Conditional Cash Transfer Program - The Nicaraguan Red de Protección Social”, Research Report 141, International Food Policy Research Institute, Washington DC.

Morris, S.S.; Olinto, P.; Flores, R.; Nilson, E.A., and Figueiró, A.C. (2004), "Conditional Cash Transfers Are Associated with a Small Reduction in the Weight Gain of Preschool Children in Northeast Brazil", *Journal of Nutrition* 134 (9), 2336–2341.

Oliveira, A. M. H. C. de; Andrade, M. V.; Resende, A. C. C.; Rodrigues, C. G.; Souza, L. R.; Ribas, R. P. (2007) "First Results of a Preliminary Evaluation of the Bolsa Família Program" in Evaluation of MDS Policies and Programs – Results Volume 2 – Bolsa Família and Social Assistance Secretariat for Evaluation and Information Management Ministry of Social Development and the Fight Against Hunger. Brasília/DF, 2007.

Olsson, U. (1979), "Maximum Likelihood Estimation of the Polychoric Correlation", *Psychometrika* 44 (4), 443-460.

Oosterbeek, H.; Ponce, J. and Schady, N. (2008. "The Impact of Unconditional Cash Transfers on School Enrollment: Evidence from Ecuador", Policy Research Working Paper 4645, World Bank, Washington DC.

Ponce, J. and Bedi, A.S. (2008), "The Impact of a Cash Transfer Program on Cognitive Achievement - The Bono de Desarrollo Humano of Ecuador", Discussion Paper 3658, Institute for the Study of Labor, Bonn, Germany.

Rawlings, L.B. and Rubio, G.M. (2005), "Evaluating the Impact of Conditional Cash Transfer Programs", *The World Bank Research Observer* 20 (1), 29-55.

Rosenbaum, P.R. and Rubin, D.B. (1983), "The Central Role of the Propensity Score in Observational Studies for Causal Effects", *Biometrika* 70 (1), 41-55.

Rubin, D.B. and Thomas, N. (1996), "Matching Using Estimated Propensity Scores - Relating Theory to Practice", *Biometrics* 52 (1), 249-264.

Schady, N. and Araujo, M.C. (2008), "Cash Transfers, Conditions, and School Enrollment in Ecuador", *Economía* 8 (2), 43–70.

Schultz, T.P. (2004). "School Subsidies for the Poor - Evaluating the Mexican PROGRESA Poverty Program", *Journal of Development Economics* 74 (1), 199–250.

Smith, J. and Todd, P. (2005), "Does Matching Overcome La Londe's Critique of Non-Experimental Estimators?", *Journal of Econometrics* 125 (1-2), 305-353.

# Appendix

## Appendix 1    Heterogeneous treatment effects 2004 (one-to-one matching)

| | Enrollment | | | | | Attendance | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **by age** | **treated** | **controls** | **Diff** | **s.e.** | **t-stat** | **treated** | **controls** | **diff** | **s.e.** | **t-stat** |
| **6-9 years** | 95.41% | 92.34% | 3.07% | 0.006 | 5.03 | 2.19 | 2.78 | -0.59 | 0.152 | -3.88 |
| **10-17 years** | 93.88% | 89.17% | 4.71% | 0.005 | 9.95 | 2.64 | 2.80 | -0.17 | 0.119 | -4.76 |
| | diff | | t-stat | | | diff | | t-stat | | |
| **younger vs. older** | -1.64% | | 3.36 *** | | | -0.42 | | -2.20 ** | | |
| **by gender** | | | | | | | | | | |
| **male** | 93.15% | 87.00% | 6.15% | 0.01 | 10.67 | 2.68 | 2.88 | -0.20 | 0.14 | -1.38 |
| **female** | 95.56% | 91.53% | 4.03% | 0.00 | 8.24 | 2.33 | 2.80 | -0.47 | 0.13 | -3.58 |
| | diff | | t-stat | | | diff | | t-stat | | |
| **Male-Female** | 2.12% | | 2.81 *** | | | 0.27 | | 1.40 | | |
| **by area** | | | | | | | | | | |
| **rural** | 93.47% | 86.37% | 7.10% | 0.01 | 6.97 | 2.33 | 2.83 | -0.51 | 0.23 | -2.21 |
| **urban** | 94.58% | 89.99% | 4.59% | 0.00 | 10.48 | 2.56 | 2.72 | -0.16 | 0.11 | -1.50 |
| | diff | | t-stat | | | diff | | t-stat | | |
| **rural-urban** | 2.51% | | 2.26 ** | | | -0.34 | | -1.36 | | |
| **by region** | | | | | | | | | | |
| **North** | 93.85% | 87.89% | 5.96% | 0.010 | 5.96 | 1.87 | 2.00 | -0.12 | 0.148 | -0.84 |
| **North-East** | 95.02% | 89.95% | 5.07% | 0.006 | 8.48 | 3.00 | 3.42 | -0.43 | 0.191 | -2.24 |
| **South-East** | 94.43% | 90.03% | 4.40% | 0.008 | 5.47 | 2.07 | 2.61 | -0.53 | 0.211 | -2.53 |
| **South** | 92.15% | 90.35% | 1.80% | 0.011 | 1.58 | 2.46 | 2.45 | 0.00 | 0.193 | 0.02 |
| **Midwest** | 94.23% | 90.92% | 3.31% | 0.012 | 2.79 | 1.87 | 2.13 | -0.26 | 0.194 | -1.36 |
| | diff | | t-stat | | | diff | | t-stat | | |
| **North-South** | 4.16% | | 2.74 | | | -0.30 | | - 0.52 | | |
| **North East - South** | 3.27% | | 2.53 | | | -0.43 | | - 1.59 | | |
| **South East - South** | 2.60% | | 1.86 | | | -0.54 | | - 1.88 ** | | |
| **Midwest - South** | 1.50% | | 0.91 | | | -0.27 | | - 0.98 | | |

## Appendix 2 Lists of Variables for the asset index

| | | | |
|---|---|---|---|
| illumination | | phone | |
| **0** | other source | 0 | no phone |
| **1** | oil, kerosene or gas (bottle) | 1 | fixed line |
| **2** | electric | 2 | only mobile phone |
| house material | | 3 | fixed line and mobile phone |
| **0** | straw | **heat kitchen** | |
| **1** | salvage wood | 0 | other combustible |
| **2** | Loam | 1 | electricity |
| **3** | wood | 2 | coal |
| **4** | brickwork | 3 | firewood |
| | | 4 | canalized gas |
| Ceiling material | | 5 | bottle gas |
| **5** | roof tile | **radio** | |
| **4** | concrete slab | 0 | no |
| **3** | wood | 1 | yes |
| **2** | zinc | **TV** | |
| **1** | salvage wood | 0 | no TV |
| **0** | straw | 1 | TV black and white |
| property condition | | 2 | color TV |
| **0** | other condition | **fridge** | |
| **1** | ceded (other) | 0 | no fridge |
| **2** | ceded by employer | 1 | yes (1 compartment) |
| **3** | rented | 2 | yes (2 compartments) |
| **4** | own property (financed) | 3 | no fridge but freezer |
| **5** | own property | 4 | freezer and fridge (1 compartment) |
| water source | | 5 | freezer and fridge (2 compartments) |
| **0** | no piped water | **herd** | |
| **1** | other source | 0 | no |
| **2** | fount or spring | 1 | 1 hot plate |
| **3** | piped network water distribution | 2 | 2 hot plates |
| nr of rooms | 1 to 30 | **washing machine** | |
| sanitation condition | | 0 | no |
| **0** | no bathroom | 1 | yes |
| **1** | other form | **computer** | |
| **2** | direct to river or lake | 0 | no computer |
| **3** | digging | 1 | computer without internet |
| **4** | rudimental fosse | 2 | computer with internet |
| **5** | cesspool (not connected to sewage) | | |
| **6** | cesspool (connected to sewage) | | |
| **7** | sewage | | |

Source: PNAD 2006

# Appendix 3 Balancing test – complete sample

| | | Mean | | | | t-test | |
|---|---|---|---|---|---|---|---|
| Variable | Sample | Treated | Control | %bias | % reduct\|bias\| | t | p>\|t\| |
| | | | | | | | |
| asset_index | Unmatched | -1.1329 | 0.11148 | -92.1 | | -128.86 | 0.000 |
| | Matched | -1.1338 | -1.2199 | 6.4 | 93.1 | 7.24 | 0.000 |
| indian | Unmatched | 0.00296 | 0.00287 | 0.2 | | 0.23 | 0.821 |
| | Matched | 0.0029 | 0.00384 | -1.7 | -980.6 | -1.97 | 0.048 |
| black | Unmatched | 0.07106 | 0.06082 | 4.1 | | 5.88 | 0.000 |
| | Matched | 0.07116 | 0.0734 | -0.9 | 78.2 | -1.06 | 0.291 |
| asian | Unmatched | 0.0015 | 0.00373 | -4.4 | | -5.79 | 0.000 |
| | Matched | 0.0015 | 0.00234 | -1.6 | 62.7 | -2.33 | 0.020 |
| parda | Unmatched | 0.64173 | 0.4613 | 36.9 | | 51.62 | 0.000 |
| | Matched | 0.64169 | 0.64746 | -1.2 | 96.8 | -1.48 | 0.140 |
| age | Unmatched | 11.327 | 11.598 | -8.0 | | -11.16 | 0.000 |
| | Matched | 11.325 | 11.257 | 2.0 | 74.7 | 2.48 | 0.013 |
| birth_registered | Unmatched | 0.9989 | 0.99797 | 2.4 | | 3.18 | 0.001 |
| | Matched | 0.99893 | 0.99743 | 3.8 | -61.0 | 4.31 | 0.000 |
| UF1 | Unmatched | 0.0151 | 0.01999 | -3.7 | | -5.14 | 0.000 |
| | Matched | 0.01515 | 0.01745 | -1.8 | 53.0 | -2.23 | 0.026 |
| UF2 | Unmatched | 0.01925 | 0.01265 | 5.3 | | 7.68 | 0.000 |
| | Matched | 0.01912 | 0.02562 | -5.2 | 1.6 | -5.39 | 0.000 |
| UF3 | Unmatched | 0.02644 | 0.02458 | 1.2 | | 1.66 | 0.096 |
| | Matched | 0.02626 | 0.03306 | -4.3 | -267.6 | -4.91 | 0.000 |
| UF4 | Unmatched | 0.00895 | 0.00542 | 4.2 | | 6.12 | 0.000 |
| | Matched | 0.00897 | 0.00801 | 1.1 | 72.6 | 1.29 | 0.197 |
| UF5 | Unmatched | 0.0675 | 0.06371 | 1.5 | | 2.17 | 0.030 |
| | Matched | 0.06739 | 0.08004 | -5.1 | -233.0 | -5.93 | 0.000 |
| UF6 | Unmatched | 0.00316 | 0.01598 | -13.2 | | -16.88 | 0.000 |
| | Matched | 0.00317 | 0.0036 | -0.4 | 96.6 | -0.91 | 0.361 |
| UF7 | Unmatched | 0.01962 | 0.01471 | 3.8 | | 5.45 | 0.000 |
| | Matched | 0.01968 | 0.02108 | -1.1 | 71.5 | -1.21 | 0.225 |
| UF8 | Unmatched | 0.03977 | 0.01387 | 16.1 | | 24.57 | 0.000 |
| | Matched | 0.03973 | 0.04284 | -1.9 | 88.0 | -1.91 | 0.056 |
| UF9 | Unmatched | 0.02996 | 0.00963 | 14.6 | | 22.48 | 0.000 |
| | Matched | 0.02996 | 0.02919 | 0.6 | 96.2 | 0.55 | 0.579 |
| UF10 | Unmatched | 0.10761 | 0.04563 | 23.5 | | 35.2 | 0.000 |
| | Matched | 0.10746 | 0.10042 | 2.7 | 88.6 | 2.82 | 0.005 |
| UF11 | Unmatched | 0.02454 | 0.01277 | 8.7 | | 12.92 | 0.000 |
| | Matched | 0.02462 | 0.02332 | 1.0 | 89.0 | 1.04 | 0.298 |
| UF12 | Unmatched | 0.03708 | 0.01184 | 16.4 | | 25.20 | 0.000 |
| | Matched | 0.03713 | 0.02889 | 5.4 | 67.4 | 5.65 | 0.000 |
| UF13 | Unmatched | 0.09098 | 0.05235 | 15.0 | | 22.04 | 0.000 |
| | Matched | 0.09105 | 0.09038 | 0.3 | 98.3 | 0.28 | 0.776 |
| UF14 | Unmatched | 0.03182 | 0.01246 | 13.2 | | 19.99 | 0.000 |
| | Matched | 0.03189 | 0.02903 | 2.0 | 85.2 | 2.04 | 0.041 |
| UF15 | Unmatched | 0.01869 | 0.01306 | 4.5 | | 6.53 | 0.000 |
| | Matched | 0.01858 | 0.01732 | 1.0 | 77.5 | 1.17 | 0.242 |
| UF16 | Unmatched | 0.13621 | 0.07446 | 20.2 | | 29.72 | 0.000 |
| | Matched | 0.13655 | 0.12734 | 3.0 | 85.1 | 3.33 | 0.001 |
| UF17 | Unmatched | 0.09517 | 0.08659 | 3.0 | | 4.23 | 0.000 |
| | Matched | 0.09528 | 0.08994 | 1.9 | 37.8 | 2.25 | 0.024 |
| UF18 | Unmatched | 0.01816 | 0.01846 | -0.2 | | -0.32 | 0.750 |
| | Matched | 0.01812 | 0.01695 | 0.9 | -284.4 | 1.09 | 0.276 |
| UF19 | Unmatched | 0.02215 | 0.06961 | -22.8 | | -29.86 | 0.000 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | Matched | 0.02222 | 0.02072 | 0.7 | 96.8 | 1.27 | 0.205 |
| UF21 | Unmatched | 0.02903 | 0.05609 | -13.4 | | -18.06 | 0.000 |
| | Matched | 0.02882 | 0.02809 | 0.4 | 97.3 | 0.54 | 0.589 |
| UF22 | Unmatched | 0.00868 | 0.03212 | -16.6 | | -21.56 | 0.000 |
| | Matched | 0.00871 | 0.00724 | 1.0 | 93.7 | 2.02 | 0.043 |
| UF23 | Unmatched | 0.0414 | 0.07642 | -14.9 | | -20.12 | 0.000 |
| | Matched | 0.0413 | 0.03913 | 0.9 | 93.8 | 1.35 | 0.177 |
| UF24 | Unmatched | 0.01081 | 0.02336 | -9.7 | | -12.93 | 0.000 |
| | Matched | 0.01084 | 0.01448 | -2.8 | 71.0 | -3.98 | 0.000 |
| UF25 | Unmatched | 0.01696 | 0.02627 | -6.4 | | -8.73 | 0.000 |
| | Matched | 0.01701 | 0.01932 | -1.6 | 75.3 | -2.11 | 0.035 |
| UF26 | Unmatched | 0.0257 | 0.04657 | -11.2 | | -15.11 | 0.000 |
| | Matched | 0.02569 | 0.02923 | -1.9 | 83.1 | -2.65 | 0.008 |
| UF27 | Unmatched | 0.01031 | 0.03578 | -17.0 | | -22.15 | 0.000 |
| | Matched | 0.01028 | 0.01171 | -1.0 | 94.4 | -1.68 | 0.092 |
| rural1 | Unmatched | 0.00486 | 0.00554 | -1.0 | | -1.33 | 0.185 |
| | Matched | 0.00487 | 0.00597 | -1.5 | -60.8 | -1.84 | 0.066 |
| rural2 | Unmatched | 0.04037 | 0.01387 | 16.4 | | 25.03 | 0.000 |
| | Matched | 0.0401 | 0.0375 | 1.6 | 90.2 | 1.65 | 0.099 |
| rural3 | Unmatched | 0.00037 | 0.00081 | -1.8 | | -2.43 | 0.015 |
| | Matched | 0.00037 | 0.00047 | -0.4 | 77.4 | -0.6 | 0.548 |
| rural4 | Unmatched | 0.00113 | 0.00041 | 2.6 | | 3.94 | 0.000 |
| | Matched | 0.00113 | 0.0008 | 1.2 | 53.5 | 1.31 | 0.189 |
| rural5 | Unmatched | 0.25981 | 0.10782 | 40.0 | | 59.58 | 0.000 |
| | Matched | 0.26046 | 0.25102 | 2.5 | 93.8 | 2.65 | 0.008 |
| urban2 | Unmatched | 0.01064 | 0.00994 | 0.7 | | 0.98 | 0.328 |
| | Matched | 0.01054 | 0.00897 | 1.6 | -125.0 | 1.95 | 0.051 |
| urban3 | Unmatched | 0.00492 | 0.00458 | 0.5 | | 0.71 | 0.477 |
| | Matched | 0.00487 | 0.00417 | 1.0 | -103.0 | 1.28 | 0.201 |
| favela | Unmatched | 0.05191 | 0.04736 | 2.1 | | 2.97 | 0.003 |
| | Matched | 0.05191 | 0.05648 | -2.1 | -0.6 | -2.47 | 0.013 |
| aldeia | Unmatched | 0.00096 | 0.00019 | 3.2 | | 5.13 | 0.000 |
| | Matched | 0.00097 | 0.0005 | 1.9 | 39.7 | 2.11 | 0.035 |
| head_age | Unmatched | 42.965 | 43.423 | -4.0 | | -5.55 | 0.000 |
| | Matched | 42.987 | 42.741 | 2.1 | 46.4 | 2.63 | 0.008 |
| head_age2 | Unmatched | 1974 | 2023.7 | -4.4 | | -6.15 | 0.000 |
| | Matched | 1975.3 | 1959.7 | 1.4 | 68.7 | 1.72 | 0.086 |
| IT_headage_Z | Unmatched | 0.44849 | 0.44805 | 0.0 | | 0.01 | 0.989 |
| | Matched | 0.44725 | 0.58511 | -3.2 | -31324.7 | -3.66 | 0.000 |
| head_lite | Unmatched | 0.70195 | 0.894 | -49.3 | | -73.92 | 0.000 |
| | Matched | 0.70121 | 0.71118 | -2.6 | 94.8 | -2.68 | 0.007 |
| head_educ | Unmatched | 4.6334 | 7.8692 | -82.4 | | -111.53 | 0.000 |
| | Matched | 4.6229 | 4.5803 | 1.1 | 98.7 | 1.58 | 0.115 |
| head_educ2 | Unmatched | 32.979 | 81.28 | -80.8 | | -105.47 | 0.000 |
| | Matched | 32.842 | 31.42 | 2.4 | 97.1 | 4.31 | 0.000 |
| head_work_~ | Unmatched | 0.32918 | 0.11677 | 52.8 | | 79.00 | 0.000 |
| | Matched | 0.32972 | 0.3029 | 6.7 | 87.4 | 7.06 | 0.000 |
| head~d_other | Unmatched | 0.00685 | 0.01178 | -5.1 | | -6.95 | 0.000 |
| | Matched | 0.00687 | 0.00674 | 0.1 | 97.3 | 0.20 | 0.842 |
| head_work_~f | Unmatched | 0.09018 | 0.12564 | -11.4 | | -15.76 | 0.000 |
| | Matched | 0.09014 | 0.09231 | -0.7 | 93.9 | -0.92 | 0.357 |
| head_work_~n | Unmatched | 0.10841 | 0.08704 | 7.2 | | 10.30 | 0.000 |
| | Matched | 0.10819 | 0.11453 | -2.1 | 70.3 | -2.47 | 0.014 |
| head_wor~rce | Unmatched | 0.09886 | 0.15635 | -17.3 | | -23.63 | 0.000 |
| | Matched | 0.09859 | 0.10512 | -2.0 | 88.6 | -2.65 | 0.008 |
| head_work~od | Unmatched | 0.02753 | 0.03271 | -3.0 | | -4.20 | 0.000 |
| | Matched | 0.02749 | 0.02696 | 0.3 | 89.7 | 0.40 | 0.688 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| head_work_~t | Unmatched | 0.03059 | 0.06266 | -15.3 | | -20.43 | 0.000 |
| | Matched | 0.03056 | 0.02859 | 0.9 | 93.9 | 1.42 | 0.155 |
| head_wor~lic | Unmatched | 0.02451 | 0.06374 | -19.2 | | -25.35 | 0.000 |
| | Matched | 0.02459 | 0.02179 | 1.4 | 92.9 | 2.28 | 0.023 |
| h_service | Unmatched | 0.02208 | 0.05457 | -17.0 | | -22.49 | 0.000 |
| | Matched | 0.02165 | 0.02139 | 0.1 | 99.2 | 0.23 | 0.822 |
| head_wor~tic | Unmatched | 0.07356 | 0.05211 | 8.8 | | 12.78 | 0.000 |
| | Matched | 0.0733 | 0.08721 | -5.7 | 35.1 | -6.27 | 0.000 |
| he~r_service | Unmatched | 0.02437 | 0.03224 | -4.7 | | -6.54 | 0.000 |
| | Matched | 0.02422 | 0.02419 | 0.0 | 99.6 | 0.03 | 0.979 |
| head~k_other | Unmatched | 0.02853 | 0.06624 | -17.8 | | -23.69 | 0.000 |
| | Matched | 0.02826 | 0.02659 | 0.8 | 95.6 | 1.25 | 0.211 |
| head_work~ed | Unmatched | 0.00545 | 0.00249 | 4.7 | | 7.05 | 0.000 |
| | Matched | 0.00547 | 0.00581 | -0.5 | 88.7 | -0.55 | 0.585 |
| mother_alive | Unmatched | 0.985 | 0.97997 | 3.8 | | 5.28 | 0.000 |
| | Matched | 0.98499 | 0.98345 | 1.2 | 69.5 | 1.51 | 0.132 |
| mother_liveHH | Unmatched | 0.9028 | 0.85261 | 15.4 | | 21.04 | 0.000 |
| | Matched | 0.90305 | 0.89364 | 2.9 | 81.3 | 3.81 | 0.000 |
| head_female | Unmatched | 0.27557 | 0.27324 | 0.5 | | 0.74 | 0.461 |
| | Matched | 0.27367 | 0.28635 | -2.8 | -443.3 | -3.46 | 0.001 |
| nr_HHhabit | Unmatched | 5.7571 | 4.6323 | 58.3 | | 85.46 | 0.000 |
| | Matched | 5.7578 | 5.7407 | 0.9 | 98.5 | 0.95 | 0.340 |
| children | Unmatched | 3.2556 | 2.2518 | 68.5 | | 101.18 | 0.000 |
| | Matched | 3.2561 | 3.2753 | -1.3 | 98.1 | -1.36 | 0.174 |
| oldest_chi | Unmatched | 13.418 | 12.842 | 17.9 | | 24.61 | 0.000 |
| | Matched | 13.418 | 13.387 | 0.9 | 94.7 | 1.19 | 0.234 |

Source: PNAD 2006, own calculation (pstest output)