The Effects of Unconditional Cash Transfers on the Home Environment in Ecuador

Juan Carlos Palacios

MSc in Econometrics Thesis

Universiteit van Amsterdam

Thesis Supervisor: Maurice Bun, Ph.D

June 2013

Abstract: Cash transfers have been increasingly popular in Latin American countries as policies aiming poverty alleviation and long run human capital accumulation. Ecuador has also implemented such policy and much interest lies in finding out whether the policy has any positive effect on people's well being. In this thesis, I analyse if households that receive the transfer offer a better environment for children's development using a randomize experiment of the transfer. I propose different techniques of estimation ranging from OLS to propensity score matching-difference in differences and analyse the validity of each technique given the conditions of the data and the cash transfer program. I find a significant improvement in the environment for children in the treated households based on the interaction between parents and children. Nevertheless, no much improvement is found in other analysed dimensions.

Introduction

Since 1990 and especially since 2000, several countries in Latin America have implemented programs of direct transfers in cash to poor households with the aim of improving their living standards. According to Fiszbein & Schady (2009) in 2008 almost every Latin American country has implemented such policy. The size of the transfers relative to pre-transfer household consumption goes from 6% to 30% considering different transfers in Latin America (Fiszbein & Schady, 2009). Therefore, large interest has lied on evaluating whether CCT's have any positive impact on the situation of the beneficiaries.

The cash transfer analysed in this study takes place in Ecuador and is named "Bono de Desarrollo Humano (BDH)", which translates to Human Development Bonus. The program was implemented in 2003 in exchange of a previously implemented cash transfer (Bono Solidario) in order to improve the allocation of the transfers by creating means to test if the household requires the transfer. The beneficiaries were ranked according to their characteristics, which are reflected by an index called SELBEN index, and the families within the two poorest quintiles of the SELBEN index would receive the transfer (Paxson & Schady, 2010). The aim of the program was providing an amount of money to mothers in condition of poverty that can alleviate the condition of their families. The monthly transfer at the moment the data was collected was USD 15, which represented around 10% of the minimum wage at the baseline[1]. The amount was doubled to USD 30 in 2007, and it is not indexed to inflation. According to Fiszbein & Schady (2009) in 2006 the BDH represented 6% of the pre-transfer consumption.

Initially, the cash transfer was conceived to be conditional on mothers taking their children to monthly health inspections and sending them to school. As Handa & Davis (2006) point out, conditioning the cash transfer impulses the long-run accumulation of human capital and alleviates the short-run poverty. However, because of lack of enforcement the conditionality was not actually implemented.

Therefore, the BDH turned into an unconditional cash transfer. It means that it can be expected it would lose some power on the long-run accumulation of human capital, and the only way it improves people's lives is through the alleviation of poverty expected from the raise in income and its chain of effects.

There are different points of view about the efficiency of the BDH and whether it improves life conditions of the beneficiaries or actually gives them non-intentioned incentives. Fiszbein & Schady (2009) point out that within the critiques of cash transfers is that people might reduce their labour supply or feel discouraged to invest in their own human capital for better future labour opportunities.

[1] The minimum wage was USD 158.10 in 2003, USD 174.9 in 2005 and USD 233.1 in 2008.

The government of Ecuador jointly with the World Bank decided to evaluate whether the BDH would have a positive effect on the development of children. Specifically, they expected that households with access to the cash transfer would provide children better nutrition and in general a better environment so physical and cognitive development would be higher among those children. Therefore they collected the data that is used in this thesis.

The evaluations of the program using the same dataset have focused on the expected final outcome: physical and cognitive development of children (Paxson & Schady, 2005, 2010; Younger, Ponce, & Hidalgo, 2009), whereas this thesis aims to check if the chain of effects holds. Therefore, I investigate whether the BDH has any impact on the environment inside the household.

The rest of this thesis is organized as follows. In Section 1 the transmission mechanisms and the set of outcome variables are described. In Section 2 I describe the data used in this study. Then in Section 3 I introduce the econometric theory relevant for treatment analysis using experimental data. In Section 4 an overview is given of the studies done in Latin America with respect to Conditional Cash Transfers. In Section 5, I review previous studies of the effects of BDH that have been investigated. In Section 6 I present the main findings of the study. I start by testing the exogeneity of treatment and assignment to treatment, then I present the results up to the 1st follow up survey and finally I make use of the baseline and the two follow up surveys to check for cumulative effects of the transfer. It is also checked the randomness of missing observations to address the external validity of the results. Finally, in Section 7 I present the most relevant conclusions of the study.

1. Outcomes of interest

There are several variables related to the environment of the household. Some are related to the physical environment and its services and others to the familiar environment.

Specifically, the analysed variables are described in Table 1. There are five broad dimensions to be explored: Construction materials and overcrowding, sanitation services, equipment of assets, general salubrity and familiar environment. All the variables were transformed such that a positive parameter reflects an improvement.

For four of the dimensions a Principal Components Analysis (PCA) was performed based on the baseline characteristics. The first factor was then taken to be analysed as an extra outcome variable besides the separate components. Based on the first eigenvector of the PCA in the baseline, the values for the first factor were constructed for the first and second follow up surveys. Given that all the variables were transformed such that a larger value implies a better condition, a larger value of the first factor of the PCA also means a better condition. The results of the individual PCA analysis can be found in the Annex 1.

Most of the variables are dummy variables indicating if the condition is satisfied. For instance, adequacy of floor material (floor_d) is a variable that takes the value of 1 if the floor in the dwelling is considered adequate. The definitions of adequacy are presented in the Annex 1 as well as the construction of all the dependent variables.

On the other hand, all the PCA obtained variables are not dummies. Likewise the HOME score ranges from 0 to 11, where 0 means a better environment. In practice I estimate the effect on the negative of the HOME so -11 means the worse possible environment and 0 the best one given the scale. The MacArthur ladders ranges from 1 to 9 and the PSS score ranges from 0 to 16, where 16 means higher perceived stress. In practice I estimate the effect on the negative PSS score. A description of the construction of these measures is in Annex 1.

The HOME score adjusted by Caldwell, B. M., & Bradley, R. H. (2003) is a scale that measures the interaction between children and their environment at their home in different dimensions for different age groups. The complete scale includes an analysis of emotional and verbal responsivity of the primary caregiver, avoidance of restriction and punishment, organisation of the physical and temporal environment, provision of appropriate play materials, parental involvement with the child and opportunities for variety in daily stimulation (Totsika & Sylva, 2004) for children between 0 and 3 years old. It has been related to cognitive development (Bradley et al., 1989; Johnson et al., 1993; Totsika & Sylva, 2004), although their findings are in general about correlation. Even after correcting for socioeconomic status, the HOME environment score has a high predictive power on different intelligence measures.

The scale used in the survey is a short form of the complete scale, which includes 11 questions described in Annex 1 about Responsivity and Punitiveness. Although the full scale has been successfully related to children development, Ferron, Ng'Andu, & Garrett (1994) found the short form scales not necessarily measure what the complete scale does.

The PSS (Perceived Stress Scale) scale is a free to use scale first proposed and used by Cohen (1983). It consists of 10 questions asking about how stressed people feel about certain events in the month prior to the survey. From the original ten questions, in the survey were collected four questions suggested as a short form by Cohen (1983). The scale has been validated and then often used as a valid tool to measure stress (Al kalaldeh & Abu Shosha, 2012).

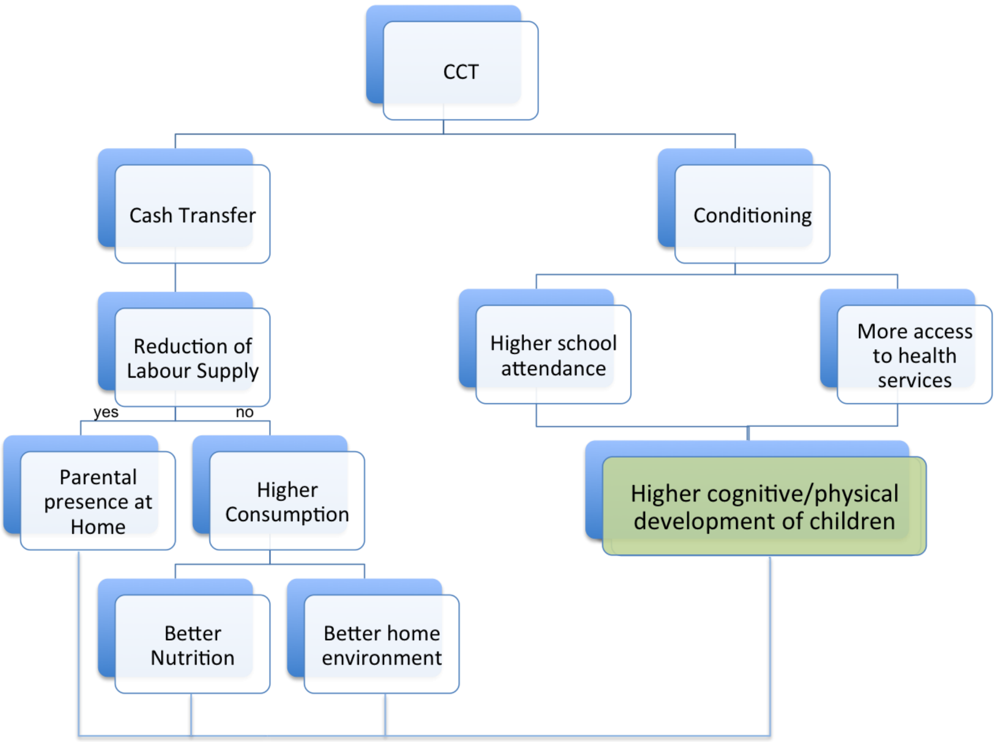
The MacArthur ladders are cards with a ladder drawn on them given to the mothers that are interviewed to indicate from 1 to 9 where she thinks her household is with respect to her community and the country. It is a way of assessing the Social Economic Status (SES). It has been used in several studies and it has been found to be from moderate to highly reliable to assess the SES (Giatti, Camelo, Rodrigues, & Barreto, 2012).

1.1 Transmission Mechanism

As described in the introduction, in the complete framework of a CCT the cash and the conditioning are expected to have positive effects on the child's development.

The expected chain of effects can be described as in Figure 1. CCT's are expected to affect and improve people's lives through an expansion of available income and via enforcing school attendance and health check-ups (Fiszbein & Schady, 2009; Handa & Davis, 2006; Villatoro, 2005).

Figure 1



The concerns about the policy usually lie on the side of the cash transfer and the incentives it gives to people. Especially politicians argue the cash transfer might create an income effect that discourages people to work. In such a case, if the lost income equals the cash transfer, the only way a household can still be benefitted from the transfer is via the parental presence at home. Felfe & Hsin (2012) found that more hours of work of the mother are related to less development of children. The transmission mechanisms are thought to be that child-mother interaction *facilitate the transmission of knowledge and skills* and/or give children the *sense of security and attachment* (Felfe & Hsin, 2012). Waldfogel (2007) makes a review of studies that investigate the relationship between parental presence and child development. Parental presence seems to be critical in many dimensions of child development with special emphasis during the first year of age.

If the cash transfer does not discourage labour supply, the household should benefit from a higher income. It's essential then to know how the new income is spent. Fiszbein & Schady (2009) point out the transfer is usually given to the mother, which is the case of the BDH, since it is expected she will spend it thinking more of her children.

If the new income is well spend, then it is expected that the availability of food and the general environment inside the house would improve. Food availability has a direct impact on physical and cognitive development via the intake of necessary micronutrients. The environment inside the house also has a relationship with the child's development. For example, the quality of housing has been pointed out to have a direct effect on development. For instance, if the floor of the dwelling is of dirt, then it is expected children to suffer more often digestive problems, which limits how many micronutrients their bodies can absorb. Leventhal & Newman (2010) review studies that analyse the relationship between quality of housing and development. They report that many studies find a positive relationship, especially related to physical development. They, however, also point out there are many studies that find no evidence of any relationship.

From the side of the conditioning, the question that mainly arises is whether the problem of low school attendance and low use of health services is a demand-side problem or is a problem of under provision or even both. This question has not been completely answered. Handa & Davis (2006) cite two studies that try to find whether supply-side interventions are more cost-efficient than demand-side interventions for school enrolment. The first study, done in Mexico, finds that demand-side policies (CCT's) are more cost-efficient. However, the second study finds that for Mozambique supply-side interventions are a better option. Anyway, both studies only consider the effect on schooling, but CCT's are also expected to have a positive impact on health provision and alleviation of poverty too, so the expected synergy obtained by the CCT can outperform supply-side interventions. Nevertheless, it is important to mention that increasing the provision of public services, especially health and education services, is always necessary, especially considering that both are considered rights under the Ecuadorian constitution. Besides, rather than being substitutes, demand-side and supply-side policies can be complementary.

Table 1

(In parenthesis the name of the Variable)

Adequacy of construction materials and overcrowding	Equipment of household	Status of general salubrity	Environment inside household
Floor (floor_d)	Fridge (DRfridge)	Isn't there excrement surrounding the house? (SANexcr_)	HOME score (Mhometot)
Roof (roof_d)	Stove (DRstove)	Isn't there garbage accumulated around the house? (SANgarb_)	MacArthur ladder in the community (MladderC)
Walls (wall_d)	Blender (DRblender)	Isn't there stagnant water around the house? (SANwatr_)	MacArthur ladder in the country (MladderE)
First Factor of PCA of construction materials (mat)	Mixer (DRmixer)	Does the house have enough ventilation? (SANvent_)	Negative PSS score[2] (Mpss)
General adequacy of construction materials[3] (matdef)	Iron (DRiron)	Is the yard or garden clean? (SANyard_)	Did the mother work in a paid job the previous week? (Mworkpay)
Not in overcrowding[4] (hac)	TV (DRcolortv)	First Factor of PCA of general salubrity (SAN)	
	Stereo (DRstereo)		
	VHS/DVD player (DRvhsdvd)		
Adequate Sanitation	Car (DRcar)		
Adequate water source (water_d)	Computer (DRcomputer)		
Adequate access to safe toilet (toilet_d)	Washing Machine (DRwasher)		
First Factor of PCA of construction materials + hac + sanitation (dwell)	First Factor of PCA of equipment (asset)		

Legend (Availability of Data)	
	On three surveys
	Baseline and 2nd Follow Up
	1st and 2nd Follow up
	Only 1st follow Up

[2] Measures the stress of the mother
[3] If all of roof_d, floor_d or wall_d are adequate, then it is considered the house has adequate construction materials
[4] A household is considered overcrowded if there are more than 3 people per bedroom.

2. Data

The government of Ecuador and the World Bank collected the data when the BDH was about to be implemented. 118 parishes[5] were selected among urban and rural parishes from 6 different provinces. They were randomly assigned to treatment and control. In the treated parishes the BDH program would be available immediately after the baseline survey and in the control parishes the program would become available after the first follow up survey. For this thesis, I only work with the parishes in the rural sector[6], where 51 parishes were chosen for the treatment group and 26 were the basis of the control group.

The baseline was collected between October 2003 and March 2004. The first follow up survey was collected between September 2005 and January 2006. After this survey, the program became available also in the control parishes. A second follow up survey was then collected between May and July 2008.

2.*1 Used data*

I use data at the household level. Initially, the dataset had 1725 different households. However, I only preserve households with complete information in all the dependent and independent variables.

First, 107 households were not found in the first follow up survey and 61 were lost on the second one, which implies an attrition of 9.7% up to the second follow up. Then, it was checked if the remaining observations had complete information on all the relevant variables in order to have the same group of analysis across all the dimensions to be analysed. On some of the variables a process of imputation was performed. Specifically, the mother's years of education were imputed based on the information of the other points in time if missing. It was checked that this variable does not vary much across time, and therefore the imputation was straightforward. Also, the variable of the HOME score for the first follow up survey was corrected for children not being awake during the interview. To collect the data on the HOME score, it is required at least one child to be awake during the interview, as collection is done by direct observation on the interaction between parents and their children. On the other variables imputation was not possible, and therefore the observations were dropped from the analysis[7]. After this process, 330 observations were lost, and the final sample is 1227 households. It means at the end I can only work with 71.13% of the original number of households.

[5] Parishes are the lowest administrative division in Ecuador with population smaller than 5,000 people.

[6] I only was able to obtain the dataset with rural observations.

[7] A balancing test on missing information was also performed to check if non-response is random, which is related to the external validity of the results. Refer to section 6.5.

In the first follow up survey, it can be identified that the compliance of the assignment to treatment is not 100%. Based on the remaining observations for this study, the take up was 82.27% in the treated parishes and 2.70% in the control ones.

3. Econometric methods for treatment analysis

First of all, in treatment analysis it is relevant to consider the Potential Outcome Model introduced by Rubin (1974). In order to identify a causal relationship and the degree of causality, it is necessary to know the counterfactual reality. The researcher needs to know what would have happened to households if they would have been assigned to the other treatment (i.e. treated households were not treated and non-treated were treated). However, it is only observed the outcome given one reality, and the other potential outcome cannot be observed and therefore has to be estimated. The aim of the different techniques is constructing a counterfactual.

The potential outcomes for any household can be seen as:

$$Y(0)\ if\ the\ household\ is\ not\ treated$$

$$Y(1)\ if\ the\ household\ is\ treated$$

$$ATE = E[Y(1) - Y(0)]$$

$$ATET = E[Y(1) - Y(0)|T = 1]$$

Where T is a dummy variable that takes the value of 1 if the household is treated and 0 otherwise, ATE is the average treatment effect and ATET the average treatment effect on the treated. The latter is obtained when the validity of the results can be only interpreted in the context of the treated group and cannot be generalized.

In reality only either Y(0) or Y(1) is observed and therefore it is necessary to construct a counterfactual to being able to estimate the treatment effect.

In order to correctly identify the average treatment effect, the unconfoundedness assumption has to be made. It implies that after controlling for observables, treatment is independent of the potential outcomes (Hirano, Imbens, & Ridder, 2003).

$$T \perp (Y(0), Y(1))|X$$

In other words, it means there are not unobservable characteristics that can be related to the take up of the treatment and the outcome. X contains all control variables that can be confounders (i.e. related to the take up and the outcome). For identification of the average treatment effect (ATE), the assumption can be relaxed to mean independence $E[Y(t)|T,X] = E[Y(t)|X], t \in [0,1]$. If the interest lies only on the average treatment effect on the treated (ATET), then it is only required $E[Y(0)|T,X] = E[Y(0)|X]$ (Hirano et al., 2003).

It is also important to make a distinction between experimental and non-experimental data. It is easier to meet the unconfoundedness assumption when using experimental data, as the assignment to treatment can be better controlled by the researcher than in the non-experimental case. The following methods to be described are focused on the use of experimental data as the data I use comes from a designed social randomized experiment.

3.1 Cross sectional Analysis

The first methods to be introduced are based on cross sectional analysis and they rely heavily on exogeneity of treatment or at least on exogeneity of assignment to treatment. Given the framework of analysis of this study, I first define the following variables:

- T_i: Dummy variable if the household lives in a parish assigned to treatment
- B_{it}: Dummy variable if the household receives the BDH in time t
- Y_{it}: outcome of interest in time t (any of the variables in Table 1)
- x_{it}[8]: vector of other k relevant explanatory variables in time t
    - Mother's education
    - Second order polynomial of mother's year of birth
    - Household size disaggregated by age groups
    - Mother is the head of the household
    - Mother's marital status
    - Municipality[9] fixed effects
- t_{1t}: dummy variable if time is in the 1st follow up survey

In treatment analysis, there are several techniques that can be used. The simplest ones usually are subject to stronger assumptions and therefore it is necessary to use more refined techniques.

[8] Some of these variables can also be affected by receiving the BDH and can be endogenous, but it is only required that after controlling for these variables, the treatment is mean independent.
[9] A municipality (canton in Ecuador) is the next administrative division after parish. Inside a municipality, there can be several parishes. Parish fixed effects were not used as treatment was randomized across parishes, which meant large multicollinearity issues.

3.1.1 OLS estimation

The first and simplest strategy would be based on a cross sectional OLS regression of the outcome of interest and a dummy variable indicating if the household receives the BDH. The information used would be in the post treatment period; specifically in t1 (i.e. the year 2005). If the unconfoundedness assumption holds to use OLS, then the Average Treatment Effect (ATE) could be estimated consistently. Specifically, the following model would be estimated:

$$Y_{i1} = \alpha + \beta_{OLS} * B_{i1} + \gamma' * x_{i1} + u_{i1} \qquad (1)$$

The requirement to identify the effect of the transfer on the outcome is mean independence of B_{i1}. It means that it is required that $E[u_{i1}|B_{i1}, x_{i1}] = f(x_{i1})$, so the only requirement is that conditional on x_{i1}, u_{i1} does not depend on B_{i1}. It allows for endogeneity in x_{i1} as long as B_{i1} is exogenous. This can hold if B_{i1} is randomly assigned conditionally or unconditionally on x_{i1} (Stock & Watson, 2008).

To see this, consider[10] $E[u_{i1}|B_{i1}, x_{i1}] = \psi' x_{i1}$, so $E[Y_{i1}|B_{i1} = 1, x_{i1}] = \alpha + \beta_{OLS} + \gamma' * x_{i1} + \psi' x_{i1}$ and $E[Y_{i1}|B_{i1} = 0, x_{i1}] = \alpha + \gamma' * x_{i1} + \psi' x_{i1}$. Therefore, $E[Y_{i1}|B_{i1} = 1, x_{i1}] - E[Y_{i1}|B_{i1} = 0, x_{i1}] = \beta_{OLS}$ (Stock & Watson, 2008). Thus, the coefficients of the control variables might be biased, but the treatment effect is not.

However, if $E[u_{i1}|B_{i1}, x_{i1}] = \psi' x_{i1} + \psi_1 * B_{i1}$, then $E[Y_{i1}|B_{i1} = 1, x_{i1}] - E[Y_{i1}|B_{i1} = 0, x_{i1}] = \beta_{OLS} + \psi_1$, where ψ_1 represents the bias that comes from the relationship between the treatment and the error term.

In order to analyse if the conditional mean assumption holds, some tests are proposed in Section 6.1. Specifically, two different tests of means are proposed and a probit model is estimated with the baseline values of the outcomes of interest on the likelihood of receiving the BDH. Under the conditional mean assumption, there should be no significant difference between those who receive the BDH and those who don't after controlling for the vector x_{it}.

3.1.2 IV estimation

If the conditional mean assumption does not hold, then one option to overcome the endogeneity problem is instrumenting B_{i1} with an exogenous variable that is related to receiving the BDH, but is not directly related to the outcome variable. The natural instrumental variable in treatment analysis is the treatment assignment. If the assumptions for this model hold, then it can be estimated using only cross sectional information in t1.

[10] If $E[u_{i1}|B_{i1}, x_{i1}] = 0$, it also holds as every variable is exogenous. The mean independence of B_{i1} is a weaker assumption than full exogeneity of all regressors.

In this case, the estimated effect is the Local Average Treatment Effect (LATE) (Imbens & Angrist, 1994) since the effect can only be extrapolated to those who comply with the instrumental variable. Compliers are those who receive the transfer if the assignment says so and otherwise do not. Specifically, the model to be estimated is a two stage least squares of the form (2SLS):

$$2nd\ Stage\quad Y_{i1} = \alpha + \beta_{IV} * \hat{B}_{i1} + \gamma' * x_{i1} + u_{i1} \tag{2}$$

$$\begin{gathered} 1st\ Stage\quad B_{i1} = \partial + \delta_1 * T_i + \zeta' * x_{i1} + v_{i1} \\ \hat{B}_{i1} = \partial + \widehat{\delta_1} * T_i + \hat{\zeta}' * x_{i1} \end{gathered} \tag{3}$$

The idea behind 2SLS is that the first stage is expected to retrieve only the exogenous part of B_{i1}, and then it can be used its estimation (i.e. $\hat{B}_{i1}$) as a regressor in (2). In practice it is performed in one single stage to adjust the standard errors for the randomness of $\hat{B}_{i1}$.

It can also be estimated the reduced form of the instrument on the outcome. In the literature of treatment analysis, this is known as the Intention to Treat Effect (ITT). The ITT can be interpreted as the average effect on the treated parishes in comparison to the control parishes. As the take up of the program in the treated parishes approximates 100% and the contamination in the control group approximates to 0%, the ITT and the LATE converges to the ATE as T_i converges to B_{i1}.

$$Y_{i1} = \alpha + \beta_{ITT} * T_i + \gamma' * x_{i1} + u_{i1} \tag{4}$$

Given T_i is a dummy variable the ITT and LATE estimators are closely related. In fact, the LATE can be retrieved from the ITT by dividing the ITT estimator by the result of the first stage. This is called the Wald Estimator and follows from the fact that under a dummy instrumental variable (Cameron & Trivedi, 2005):

$$\widehat{\beta_{IV}} = \frac{\widehat{\beta_{ITT}}}{\hat{\delta}}$$

Given $\hat{\delta}$ is positive, the ITT and 2SLS estimation would always have the same sign. The only difference is that the 2SLS estimator is then scaled by the estimated rate of compliance (i.e. δ).

In the case of the IV estimation[11], in matrix notation the estimated parameters can be expressed as:

$$\widehat{\theta_{IV}} = \theta + (Z'X)^{-1}Z'u \tag{5}$$

[11] Given it is just identified

Where

$$\theta = \begin{bmatrix} \alpha \\ \beta_{IV} \\ \gamma_1 \\ \vdots \\ \gamma_k \end{bmatrix} \qquad X = \begin{bmatrix} 1 & B_{11} & x_{111} & \cdots & x_{k11} \\ 1 & B_{21} & x_{121} & \cdots & x_{k21} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & B_{n1} & x_{1n1} & \cdots & x_{kn1} \end{bmatrix}$$

$$Z = \begin{bmatrix} 1 & T_1 & x_{111} & \cdots & x_{k11} \\ 1 & T_2 & x_{121} & \cdots & x_{k21} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & T_n & x_{1n1} & \cdots & x_{kn1} \end{bmatrix} \qquad u = \begin{bmatrix} u_{11} \\ u_{21} \\ \vdots \\ u_{n1} \end{bmatrix}$$

The variable B_i is expected to be endogenous and therefore T_i replaces it in Z. The conditions for this estimation to be unbiased is $E[\widehat{\theta_{IV}}] = \theta$, which requires $E[u|Z, X] = 0$, which requires $E[u|X] = 0$ (Cameron & Trivedi, 2005). The last condition says X is exogenous, which would mean that 2SLS is not required in the first place. Therefore IV estimation relies on asymptotic consistency[12], which relies on a large sample size.

Taking the probability limit of $\widehat{\theta_{IV}}$:

$$plim(\widehat{\theta_{IV}}) = \theta + plim(N^{-1}Z'X)^{-1} * plim(N^{-1}Z'u) \qquad (6)$$

The requirements for consistency are $plim(N^{-1}Z'X)^{-1} \neq 0$ and $plim(N^{-1}Z'u) = 0$. Where N is the sample size.

$plim(N^{-1}Z'u) = 0$ requires that $E[z_{i1}'u_{i1}] = 0$. Given $E[u_{i1}] = 0$, it is needed that $cov(z_{i1}, u_{i1}) = 0$ for the IV estimator to be consistent, where $z_{i1}' = [1 \quad T_i \quad x_{1i1} \quad \cdots \quad x_{ki1}]$. In other words, it requires z_{i1} to be exogenous. In order to test if z_i is exogenous, assuming all the control variables are exogenous, I test for T_i to be exogenous based on the same tests I perform to check the exogeneity of B_{i1}.

The ITT estimation is the OLS estimator of the form:

$$\widehat{\theta_{ITT}} = \theta + (Z'Z)^{-1}Z'u \qquad (7)$$

For $\widehat{\theta_{ITT}}$ to be unbiased, the conditional mean assumption should hold for T_i, which is later analysed in Section 6.1 when the exogeneity tests of T_i are presented. If T_i fails to be exogenous, both the IV and the ITT estimation would be inconsistent and biased.

[12] The 2SLS estimator is always biased

3.2 Panel Data Analysis

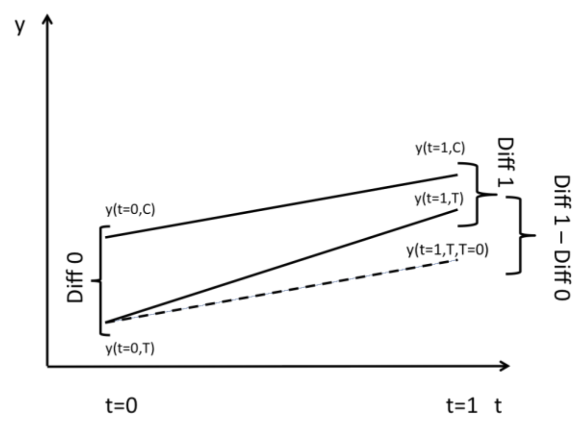
3.2.1 Difference in differences (DiD) estimator

If treatment and assignment to treatment are not exogenous, the estimated effects using cross sectional approaches would be expected to be inconsistent and biased. One possibility where the cross sectional techniques might not work is if the initial conditions of the treatment groups differ. In order to eliminate the systematic difference between the two groups, a Difference in Differences (DiD) estimator is proposed.

Based on Figure 2, the proposed techniques can be analysed. If the difference in time 0 is non-existent for the groups that do and do not receive the BDH or the groups assigned to treatment and control, then cross sectional OLS or 2SLS/ITT techniques could have been used, respectively. The cross sectional OLS estimator would give the same results as the DiD estimator under these conditions.

The DiD estimator uses the information in time 0 to correct for systematic differences between both groups. Therefore, it creates a new counterfactual: y(t=1,T,T=0) in Figure 2. It represents what is the expected outcome (y) for the treated group in time 1 if they would not have been treated. It assumes that the treated group would have had the same trend as the control group in y if the treatment had not occurred. This assumption known as parallel-trend assumption (Khandker, B. Koolwal, & Samad, 2009) implies that the systematic difference, known as unobserved heterogeneity, between both groups is not time variant and that it is the only source of bias in the cross sectional analysis, which is required for the conditional mean independence assumption to hold. Therefore, the estimator of the treatment effect is the difference between the differences in the two periods of time, hence the term difference in differences. In the example, the cross sectional analysis would have estimated a negative treatment effect (Diff 1), whereas the DiD would have found a positive one after correcting for the unobserved heterogeneity (Diff 1 – Diff 0).

Figure 2



The estimation can be performed via the following OLS regression:

$$Y_{it} = \alpha + \beta_{DiD} * B_{it} + \beta_2 * B_{i1} + \beta_3 * t_{1t} + \gamma' * x_{it} + u_{it} \tag{8}$$

Where β_2 absorbs the systematic difference between the two groups (Diff 0), β_3 absorbs the parallel time trend, and β_{DID} identifies the Average Treatment Effect on the Treated (ATET). In the literature usually B_{it} is written as an interaction term between B_{i1} and t_{1t}. Both are equal as in time 0 nobody receives the BDH.

Given the two groups start from a different point in time 0, the results cannot be generalized and the effect is the effect given the initial conditions of the treated, hence it can be only understood as the average treatment effect on the treated (ATET). It is assumed that $E[Y(0)|B,X] = E[Y(0)|X]$, where

$$X = \begin{bmatrix} 1 & B_{10} & B_{11} & 0 & x_{110} & \cdots & x_{k10} \\ 1 & B_{11} & B_{11} & 1 & x_{111} & \cdots & x_{k11} \\ 1 & B_{20} & B_{21} & 0 & x_{120} & \cdots & x_{k20} \\ 1 & B_{21} & B_{21} & 1 & x_{121} & \cdots & x_{k21} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & B_{n0} & B_{n1} & 0 & x_{1n0} & \cdots & x_{kn0} \\ 1 & B_{n1} & B_{n1} & 1 & x_{1n1} & \cdots & x_{kn1} \end{bmatrix}$$

However, $E[Y(B)|B,X] = E[Y(B)|X], B \in [0,1]$ cannot be assumed.

The DiD estimator corrects the cross sectional OLS for group unobserved heterogeneity, which is assumed to be time invariant and is eliminated by the inclusion of B_{i1}. In the cross sectional analysis, if there were intrinsic differences between groups, it is not taken into account and therefore such difference goes into the error term. If those intrinsic differences are related to receiving the BDH or to being treated, the OLS and IV estimation are invalidated because of lack of exogeneity.

In order to perform a DiD estimation it is required to have information on both periods. However, for some of the variables analysed on this thesis, the baseline information is not available.

If the cross sectional bias is non-existent (no systematic differences between treated and non-treated), then β_2 would not be significantly different from zero and the OLS and DiD estimators would converge to the same value. For unbiasedness of the DiD estimator, it is assumed the only source of bias is the group unobserved heterogeneity, which is time invariant and does not change within households of the same group[13]. If other sources of bias were present or the heterogeneity between both groups change over time, then the DiD estimation would also be biased. The identifying assumption of the DiD estimation is the

[13] And if it varies within households of the same group, it is expected to be purely random and not related to the take up of the transfer.

common trend assumption, which cannot be tested if information pre-baseline is not available. Later in Section 6.4, I proposed an informal test of the common trend assumption using the three surveys.

3.2.2 Fixed Effects (FE) estimator

It is possible that there exists not only group unobserved heterogeneity, but also individual unobserved heterogeneity, which cannot be taken care by the DiD. Consider the following data generating process:

$$Y_{it} = \alpha_i + \beta_{FE} * B_{it} + \beta_2 * B_{i1} + \delta * t_{1t} + \gamma' * x_{it} + u_{it} \tag{9}$$

where α_i reflects the time invariant individual fixed effect or individual unobserved heterogeneity and β_2 keeps reflecting the group fixed effect. In the case of the DiD estimator, a common intercept is estimated, which yield:

$$Y_{it} = \alpha + \beta_{DID} * B_{it} + \beta_2 * B_{i1} + \delta * t_{1t} + \gamma' * x_{it} + v_{it} \tag{10}$$

$$v_{it} = \alpha_i - \alpha + u_{it} \tag{11}$$

where u_{it} represents a purely random error term. If DiD is used and α_i is correlated with the taking up of the program, then there would be some endogeneity left in the residual. The fixed effects (FE) model allows households to have a time invariant individual effect, which is eliminated by subtracting the individual mean. This estimator is known as the within estimator (Cameron & Trivedi, 2005). If there are two periods in time:

$$Y_{i0} = \alpha_i + \beta_{FE} * B_{i0} + \beta_2 * B_{i1} + \gamma' * x_{i0} + u_{i0} \tag{12}$$

$$Y_{i1} = \alpha_i + \beta_{FE} * B_{i1} + \beta_2 * B_{i1} + \delta + \gamma' * x_{i1} + u_{i1} \tag{13}$$

Define, the individual mean

$$\bar{Y}_i = \frac{Y_{i0} + Y_{i1}}{2} = \alpha_i + \beta_{FE} * \bar{B}_i + \beta_2 * B_{i1} + \delta * \frac{1}{2} + \gamma' \bar{x}_i * + \bar{u}_i \tag{14}$$

By differencing the equations (12) and (13) with respect to (14), the individual and the group effects are dropped.

$$Y_{it} - \bar{Y}_i = \beta_{FE} * (B_{it} - \bar{B}_i) + \delta * (t_t - 1/2) + \gamma' * (x_{it} - \bar{x}_i) + (u_{it} - \bar{u}_i) \tag{15}$$

If the endogeneity problem came from the individual and/or group unobserved heterogeneity, then the ATET can be consistently estimated applying OLS on (15). To see this, consider the Least Squares Estimation with Dummy Variables (LSDV), which is equal to the within estimator (Cameron & Trivedi, 2005). The LSDV estimation performs an OLS estimation in (9) by including N individual dummy variables. Therefore, the matrix X becomes:

$$X = \begin{bmatrix} 1 & 0 & \cdots & 0 & B_{10} & B_{11} & 0 & x_{110} & \cdots & x_{k10} \\ 1 & 0 & \cdots & 0 & B_{11} & B_{11} & 1 & x_{111} & \cdots & x_{k11} \\ 0 & 1 & \cdots & 0 & B_{20} & B_{21} & 0 & x_{120} & \cdots & x_{k20} \\ 0 & 1 & \cdots & 0 & B_{21} & B_{21} & 1 & x_{121} & \cdots & x_{k21} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & B_{n0} & B_{n1} & 0 & x_{1n0} & \cdots & x_{kn0} \\ 0 & 0 & \cdots & 1 & B_{n1} & B_{n1} & 1 & x_{1n1} & \cdots & x_{kn1} \end{bmatrix}$$

Therefore, if either group or individual fixed effects were the source of bias in the cross sectional analysis, by including them in the matrix X, $E[Y(0)|T, X] = E[Y(0)|X]$ would hold, and the ATET can be consistently estimated.

It is important to mention the FE model is an improvement over the DiD model since it takes into account the individual effects, whereas the DiD estimator only gets rid of the systematic difference between groups as a whole. The DiD and FE estimators are equal to each other if within groups of treatment (i.e. treated and control) there are no differences among households, so the only unobserved heterogeneity is that of the group, which is controlled by both methods.

The assumption to consistently estimate the treatment effect is that under the FE and DiD models unobserved heterogeneity, either individual and/or group, is the only source of bias and it is time invariant. One drawback of the FE model is that the effect of time invariant variables cannot be estimated, i.e. in this study: mother's year of birth and municipality fixed effects. This does not mean the estimation might suffer from omitted variable bias as any fixed effect is differentiated out; this drawback would be relevant if the aim of the paper would be estimating the effect of a time invariant variable. The treatment variable is time variant as in the baseline nobody receives the transfer, and the status of some households change for the first follow up survey and also for the second follow up survey. The latter would be relevant in a posterior section of this study.

Given such assumptions, the FE estimator would be of the form:

$$\widehat{\theta_{FE}} = \theta + (X'X)^{-1}X'u \tag{16}$$

Where $E[\widehat{\theta_{FE}}] = \theta$ given the correlation between u and X only came from the unobserved heterogeneity and now that it has been dropped out, u is a purely random shock not related to X. There can even be endogeneity in the control variables, as long as the conditional mean independence assumption holds. It means it can be accepted that $E[u_{it}|B_{it}, x_{it}, B_{1t}, \alpha_i, t_t] = f(x_{it}, B_{1t}, \alpha_i, t_t)$ as long as the conditional expectation of u_{it} does not depend on B_{it}. The proof of it follows from the proof in section 3.1.1.

3.2.3 Propensity Score Matching Difference in Differences (PSM-DiD) estimator

Finally, I follow the recommendation on Khandker et al. (2009) to use a mixture of the Propensity Score Matching (PSM) and the Difference in Differences (DiD) methods. PSM methods rely on using as counterfactual for each treated observation, a weighted sum of the output of the control observations. The weights are constructed based on how similar the control observations are with respect to the analysed treated observation. In order to analyse how similar controls are to treated observations, a propensity score is estimated via a probit or logit model based on baseline characteristics. It estimates the probability of being treated given a set of covariates.

Different PSM methods use different functions to construct the weights. For instance, nearest neighbour PSM gives a weight of 1 to the control observation that is closest to the treated one based on their propensity scores and the other control observations receive a weight of 0. Other functions give different weights; for this study a kernel matching procedure is used, which gives every observation in the control group under the common support a weight based on their propensity scores. It gives higher weights to closer observations and the sum of weights is 1.

The common support is constituted by the observations of the control and treated groups that overlap based on the propensity score. For example, if the propensity score for the control group goes from 0.2 to 0.8, and for the treated from 0.4 to 0.99, then the common support are all observations with propensity score from 0.4 to 0.8. A small common support leads to biased estimates. PSM methods rely on the fact that only differences in observables explain the take up or not of the program. They rely again on the unconfoundedness assumption and that matching on the propensity score is enough to control for all the covariates, which follows from the next proof from Dehejia & Wahba (1998).

If $E[Y_i(0)|D_i, X_i] = E[Y_i(0)|X_i]$ holds, then the ATET can be identified as

$$ATET = E[Y_i(1)|D_i = 1] - E[Y_i(0)|D_i = 1]$$

$$ATET = E_X[\{E[Y_i(1)|X_i, D_i = 1] - E[Y_i(0)|X_i, D_i = 1]\}|D_i = 1]$$

$$ATET = E_X[\{E[Y_i(1)|X_i, D_i = 1] - E[Y_i(0)|X_i, D_i = 0]\}|D_i = 1]$$

It means that after conditioning on X, the output without treatment would have been the same for treated and non-treated. In the case of the DiD-PSM estimator, it means that after conditioning on X, the change in Y would have been the same for treated and non-treated observations, which is another way of looking at the common trend assumption.

Furthermore, Dehejia & Wahba (1998) prove that if $E[Y_i(0)|D_i, X_i] = E[Y_i(0)|X_i]$, then $E[Y_i(0)|D_i, p(X_i)] = E[Y_i(0)|p(X_i)]$. Define $p(X_i)$ as the propensity score equal to the probability of receiving treatment given X_i.

$$\Pr(T_i = 1|X_i) = E(T_i|X_i) = p(X_i)$$

Then,

$$\begin{aligned} E(T_i|Y_i(1), Y_i(0), p(X_i)) &= E_X\{E(T_i|Y_i(1), Y_i(0), X_i)|Y_i(1), Y_i(0), p(X_i)\} \\ &= E_X\{E(T_i|X_i)|Y_i(1), Y_i(0), p(X_i)\} \\ &= E_X\{p(X_i)|Y_i(1), Y_i(0), p(X_i)\} \\ &= p(X_i) \end{aligned}$$

And from the previous proofs, it can be concluded that conditioning on the propensity score ($p(X_i)$) is as good as conditioning on the set of covariates in X (Dehejia & Wahba, 1998).

Specifically for this study, the propensity score is estimated as the probability of receiving the transfer via a probit model based on baseline characteristics. The observations for the probit model were those households in the treated parishes only, as the program was not available for the control parishes. Based on the parameters of that model, the propensity score was estimated for the entire sample.

Later, the effect for the observation i is estimated by $DiD_i = (y_{i1}^T - y_{i0}^T) - \sum_j \omega\,(i,j)\big(y_{j1}^C - y_{j0}^C\big)$, where y_{it}^S is the output for observation i in time t that has the status S of treatment (i.e. treated (T) or control (C)). $\omega(i,j)$ is the weight of observation j for observation i based on the propensity score. Specifically, an epanechnikov kernel function is used to estimate the weights.

$$\omega(i,j) = \frac{K\left(\frac{P_j - P_i}{a_n}\right)}{\sum_{k\in C} K\left(\frac{P_k - P_i}{a_n}\right)} \tag{17}$$

Where P_j is the propensity score of the observation j, a_n is a bandwidth parameter and K is the kernel function:

$$K(u) = \frac{3}{4}(1 - u^2) * 1_{\{|u|\leq 1\}} \tag{18}$$

Both j and k refer to observations in the control group in the common support and i belongs to the treated group (i.e. those who receive the BDH).

To obtain the average treatment effect on the treated (ATET), the average of the effects for each treated observation is estimated. The standard errors are calculated via 200 bootstrap replications[14]. Given that it is performed the first difference for each observation (i.e. $(y_{i1}^{T} - y_{i0}^{T})$), it can actually be considered a Fixed Effects-PSM method.

In principle the advantage of using a matching estimator over a regression based one is the matching estimator can be consider as a semi-parametric one. It is non-parametric on the estimation of the treatment effect, but parametric on the estimation of the propensity score via a probit model. The nature of being semi-parametric helps it avoiding a potential misspecification of the regression model (Grilli & Rampichini, 2011). However, if the regression specification is correct, then OLS estimation is more efficient (i.e. has a smaller variance) (Grilli & Rampichini, 2011).

Despite the apparent advantage of relaxing the linearity of the regression, matching methods include many sources of arbitrariness that can influence the estimation:

- The choice of the set of covariates in X.
- The matching method.
- Finally, Abadie & Imbens (2008) show that bootstrap standard errors do not give asymptotically correct standard errors for matching methods.

Given those considerations, the results of the PSM-DiD are presented as a reference, but larger attention is paid to the DiD and FE estimators. Three different estimations for the PSM-DiD are actually performed: one with the presented Kernel (shown in Table 2), one with Local Linear Regressions (LLM) and one with the closest 5 neighbours to have a set of comparable estimators.

The LLM *"estimates a nonparametric locally weighted regression of the comparison group outcome in the neighbourhood of each treated"* (Khandker et al., 2009). It changes the weighting function to:

$$\omega(i,j) = \frac{K_{ij}\sum_{k\in C} K_{ik}(P_k - P_i)^2 - [K_{ij}(P_j - P_i)]\sum_{k\in C} K_{ik}(P_k - P_i)}{\sum_{j\in C} K_{ij}\sum_{k\in C} K_{ij}\,(P_k - P_i)^2 - [\sum_{k\in C} K_{ik}(P_k - P_i)]^2} \tag{19}$$

Where K_{ij} is a kernel function of p_j with respect to p_i. The common choice is a tricube kernel function.

The 5 nearest neighbours approach gives a weight of 1/5 to each one of the 5 closest control observations based on the propensity score for every treated observation and 0 otherwise.

In order to estimate the PSM-DiD model I use the Leuven & Sianesi (2003) PSM module for Stata.

[14] It was bootstrapped jointly the probit model and the matching procedure.

4. Previous studies on Conditional Cash Transfers in Latin America

Even though CCT's have become more popular around the world, most programs and therefore most evaluations have been performed in Latin America. Fiszbein & Schady (2009) made an extensive compilation of the studies related to cash transfers, which is the base of this section.

4. 1 Impact on labour supply

The evaluation of CCT programs has focused on the effects on child and adult labour. Reducing child labour is a policy objective, whereas negative effects on adult labour are unintentional effects that policy makers want to avoid. However, both can affect the level of income/expenditure in the household and as found by many authors, a larger presence at home might affect positively the child's development.

According to the compilation made by Fiszbein & Schady (2009), most CCT programs have been found to reduce child labour and especially for older children. Oportunidades in Mexico was found to reduce child labour substantially for children aged 12-17 with a higher effect for boys using a Differences in Differences (DiD) design on quasi-experimental data[15] (Skoufias & Parker, 2001). Likewise, Edmonds & Schady (2012), using a 2SLS approach on experimental data, found for Ecuador a significant reduction of child labour, especially that related to economic activities (i.e. not household chores). Significant effects were also found in Nicaragua for the RPS program using a DiD approach on experimental data (Maluccio & Flores, 2004). For the cases of Colombia and Honduras no significant effect was found (Fiszbein & Schady, 2009).

Even though there exists major concern about the possible disincentives to work CCT programs give adults, the results seem not to confirm such idea. The main mechanism through which a CCT is expected to reduce labour supply is via an income effect such that the person feels richer and can afford leisure. Of the reviewed studies by Fiszbein & Schady (2009), only one study found significant disincentives to adult labour in Nicaragua.

Fiszbein & Schady (2009) claim that the results can be explained from different points of view. First, it is very likely that in poor households, leisure elasticity is close to 0. It means that changes in income are not followed by changes in labour supply. They also claim it is important to take into account that child labour is usually significantly reduced so the total income of the household might even go down if the loss income is higher than the transfer. Besides, the conditioning on

[15] Oportunidades was made available in different regions progressively, and therefore there were regions that at to some point were not already treated and other that were. Regions are not necessarily comparable, but the authors worked on choosing only eligible households.

schooling might imply new out of pocket expenditures that do not allow adults to reduce their labour supply. As a matter of fact, given those effects, increasing labour supply, or at least not reducing it, can be a way to maintain the pre-transfer level of consumption (Fiszbein & Schady, 2009).

In the case of Nicaragua, its CCT program represents almost 30% of the pre-transfer consumption and is therefore the most generous program. Given those conditions, one study finds a significant reduction in adult labour on men, but no effect on women (Fiszbein & Schady, 2009).

4.2 Impact on consumption

There was found in several studies positive effects on consumption ranging from 7% to 29% for different countries. Specifically it was found a positive significant increment of consumption in Brazil, Colombia, Honduras, Mexico and Nicaragua. For Ecuador, there was no significant effect found (Fiszbein & Schady, 2009). All of the estimates were obtained using a DiD approach, except for the Mexican and Brazilian case, where no data was available before the program was implemented. It was also found a short-run poverty reduction in Colombia, Honduras, Mexico and Nicaragua using again a DiD approach, except for Mexico (Fiszbein & Schady, 2009).

On long-run effects, if the transfer stopped, two studies in Mexico found that recipients of Oportunidades make larger productive investment, which would help them increase their long-run consumption permanently.

4.3 Impact on School enrolment

Considering studies for Chile, Colombia, Ecuador, Honduras, Jamaica, Mexico and Nicaragua, Fiszbein & Schady (2009) summarize that there has been found significant increases in school enrolment for all the mentioned countries. The lowest effect on regular school enrolment was 2.1% increase in Colombia for children aged 8 to 13. The largest effect was found in Nicaragua with a significant effect of 12.8% increase in enrolment for children between 7 and 13 years old.

The studies, except for Ecuador, Chile, Colombia and Jamaica, were conducted using experimental data relying and testing on the randomization process. For the case of Ecuador, the data is also experimental, but due to partial compliance an IV mixed with DiD estimation was performed. In the Colombian case, they use a PSM-DiD approach. In the two remaining cases they used Regression Discontinuity approaches. The latter relies on a drastic change in the probability of receiving the transfer when a threshold of a continuous variable is surpassed. The identifying assumption of such procedure is that in the neighbourhood of the threshold observations are highly comparable.

4.4 Impact on Health check ups

In the studies cited by Fiszbein & Schady (2009), there are countries like Colombia, Jamaica, Nicaragua and Honduras, where there was found an increase in health check ups, but also countries like Chile, Ecuador and Mexico where no effect was found. Likewise, some of the CCT had a positive effect on vaccination coverage. The estimation procedures follow the same as in the school enrolment analysis.

4.5 Impact on children development

Firstly, Fiszbein & Schady (2009) revise the effects on physical development by the height for age z-score. In the studies they review, there was found a positive effect using for Colombian children younger than 24 months olds using an PSM-DiD estimation. Likewise, an improvement on Mexican children between 12 and 36 months old and younger than 6 months old was found via a cross sectional analysis relying on the randomization of the program. Finally, Nicaraguan children younger than 60 months old also improved their growth; the estimation was obtained by a cross sectional analysis of a randomized experiment. In Ecuador, Brazil and Honduras no effect was found.

Then, Fiszbein & Schady (2009) compare two studies about cognitive development for Ecuador and Nicaragua. The Nicaraguan transfer has positive effects in many of the tests, whereas in the Ecuadorian case there is only a significant effect, with 10 per cent significance, on long-term memory for the poorest quartile. Both studies presented use data on randomized experiments, and the estimation is performed using cross sectional information.

In general, it seems CCT policies have positive impacts, but how large these effects are depend much on the size and conditionality of the program. Larger effects have been found for the Nicaraguan case, but also larger unintentional disincentives to adult labour supply given it is the most generous program.

5. Chain of effects for BDH

Even though the BDH was not conditioned until recently, and the enforcement is not yet completely implemented, much interest has lied on knowing if the BDH affects school attendance/enrolment and medical check ups.

The techniques to test if there is any effect are both experimental and non-experimental. Most of the analyses exploit non-experimental techniques given the policy is available to anyone who meets the requirements. Two datasets were collected as experimental data, although both lack perfect compliance[16].

Araujo & Schady (2008) use a randomized experiment with imperfect compliance and apply a Difference in Differences approach where the dependent variable was the change in enrolment between the first follow up survey and the baseline. The explanatory variables were a set of baseline characteristics and a dummy of receiving the BDH, which was instrumented by the randomization to treatment. They found a 10.3% increase in school enrolment for children between 6 to 13 years old. They found that the increase is especially concentrated in the fifth and eight grades, where much of the dropouts are registered. Although the conditionality at the time was not actually implemented neither was it enforced, some people in the survey thought of the transfer as conditional on sending children to school. Araujo & Schady (2008) estimated if the effect was larger for households who thought the transfer was conditional and found, under different specifications, the effect was always non-negative. For some specifications the effect was not significantly different from zero. Likewise, Edmonds & Schady (2012) found a 19% increment in school enrolment for children between 11 and 16 years old (age in the baseline; between 13 and 18 in the follow up survey). They use the same dataset as Araujo & Schady (2008) and use a cross sectional instrumental variable approach, where receiving the BDH is instrumented by the randomization of the program.

On a health evaluation, Paxson & Schady (2010), using the same dataset that is used in this thesis, found that the BDH does not increase the number of health check-ups. However, they found a significant increment in the usage of parasite treatment for the poorest quartile. They use an intention to treat approach, where assignment to treatment is the relevant independent variable, and its parameter reflects the intention to treat.

As mentioned before, there has been found a significant decrease in child labour, but there are no signs of reduction of adult labour (Edmonds & Schady, 2012). When analysing consumption, there has not been found any significant effect.

However, it is important not only to understand if the aggregate consumption is affected, but also its composition. Specifically, Schady & Rosero (2008) found that the share of food on total consumption increases for the randomly selected households receiving the BDH. This is an intention to treat analysis, as not all the selected households actually signed in for the program, and many of the not selected did. This result is found based on a non-parametric construction of Engel

[16] Perfect compliance means that all households assigned to treatment actually receives it and otherwise do not.

curves for selected and not selected households. It was then confirmed by an OLS regression of assignment to treatment on food share on consumption. A 2SLS approach failed to find any significant effect, although it can be related with the higher standard errors 2SLS gives.

On the other hand, Nabernegg (2012) using a fuzzy regression discontinuity (RD) design checks if BDH recipients spend more on "undesirable" goods such as tobacco, beer, other kinds of alcohol and cellular phone. The BDH was given based on an index (SELBEN index[17]) that ranges from 0 to 100. Usually the cut-off point is 50.65, although the probability of receiving it does not go from 1 to 0 after this threshold, which explains the use of a "fuzzy" RD. He didn't find any evidence of undesirable use of BDH. In a fuzzy RD design the consumption on these goods are regressed on receiving the BDH, and receiving the BDH is instrumented by a dummy if the household is bellow the cut-off point of 50.65. The observations used for the estimation lie in the neighbourhood of the threshold.

Lastly, on the study of the impact on development there have been several studies that investigate how development is affected by the BDH. Using the same dataset as the one in this thesis, Paxson & Schady (2010) analyse a series of physical and cognitive measures of children younger than 84 months. They estimate the intention to treat effect (ITT) by regressing each one of the z-score of the tests on being selected to receive the BDH in a seemingly unrelated regressions (SUR) system. They cannot use receiving the BDH as independent variable, as they cannot rule out its endogeneity. They found the effect concentrates among the poorest quartile. In the most complete specification, they only found an improvement in the haemoglobin level at the 1% significance level for children in the poorest quartile. They also found that in the poorest quartile there was an improvement of the mother's haemoglobin level, the self-reported position in community and the environment inside the house measured by the HOME scale.

Fernald & Hidrobo (2011) using the same dataset found that in rural areas in parishes assigned to treatment children have larger developments in language compared to the children in parishes assigned to control using an ITT approach.

Ponce & Bedi (2010) uses cross sectional data on standardized mathematics and language tests to check if there is any effect of the BDH on these measures. The data they use also includes information about student's teachers, household characteristics and whether they receive the BDH. Given the data is not experimental, then their strategy is using a fuzzy regression discontinuity design by instrumenting receiving the BDH with a dummy variable indicating whether a household has a SELBEN index smaller than 50.65 just like Nabernegg (2012). They found no evidence of any significant impact on these tests.

[17] The methodology of evaluation was updated in 2009 and the index is currently called Índice RS.

Younger, Ponce, & Hidalgo (2009) use a difference in differences approach using the same dataset as here to test if there is any significant effect on height for age z-score and haemoglobin. Their findings are that there is no effect on these measures.

Finally, Schady (2012) explores if the BDH has an effect on the health status of the mother using the same dataset to be used in this thesis. He finds, using an instrumental variables approach, that the transfer significantly increases the level of haemoglobin, which was complemented by a reduction in the prevalence of anaemia.

6. Main Results

In this section, I present the principal results obtained in the framework detailed in the former sections. I first start by testing the exogeneity of receiving the BDH and the assignment to receive the BDH. If the tests show one of them is exogenous, then cross sectional methodologies can be used, and otherwise panel approaches should be performed to estimate the effects.

6.1 Exogeneity Tests

6.1.1 Treatment

To check whether receiving the BDH is exogenous three tests were performed on the variable B_{i1}. First, a simple test is performed comparing of means of the baseline characteristics between recipients and non-recipients of the BDH. There was hardly any significant difference across both groups. Then, the same test was performed but controlling for the set of control variables in the vector x_i. The results in this case adjusted the p-values and many more variables are found to differ across both groups especially in items related to the construction materials and the assets of the household. The recipients are in general worse off than not recipients. Finally, a probit model was estimated where the dependent variable was receiving the BDH and the covariates were the set of baseline values of the control and outcome variables. Under the hypothesis that receiving the BDH is exogenous, none of the baseline characteristics should explain the likelihood of receiving the transfer. It was found the age of the mother, the mother's education, the situation of overcrowding and the assets explain significantly the likelihood of receiving the BDH. In all the tests, the standard errors were clustered at the parish level. The results can be found in the Annex 2.

Another way to see how the bias works is graphically in Figure 2 (p.13). The OLS approach only measures the difference in time 1 between treated and non-treated households. If initially treated households are worse off, the estimated effect via OLS is negatively biased and it can even have a negative sign if the treatment effect is not large enough to close the gap between treated and non-treated households as shown in the figure.

Summarizing, the fact that households in worse initial conditions are more prone to receive the BDH is not in line with the required of exogeneity of treatment. This auto selection procedure creates a negative bias in the estimator of the treatment effect, which violates the mean independence assumption.

6.1.2 Assignment to treatment

The requirement to being able to identity both the LATE and the ITT effects is that the assignment to treatment must be exogenous. The process of assignment was done randomly, but there were some loss observations because of missing information and the randomization was across 77 parishes, which can lead to some unbalance between both groups given the number of randomized units. In order to check for exogeneity, the same procedures as for treatment (B_{i1}) were performed for T_i. The first and second tests gave hardly any significant difference between the two groups. However, the probit model estimated that the mother's education, the general situation of the construction materials and the possession of some assets are significantly related to the fact of living in a treated parish. The treated group seems to be systematically worse off than the control ones. Again, the standard errors were clustered at the parish level for the three tests.

Even though many of the baseline characteristics are not significantly different between both groups, the endogeneity of the assignment to treatment cannot be ruled out for the sample I use. The results can be found in the Annex 2.

However, given that households in treated parishes seem to be worse off than households in control parishes for my sample, the covariance between the error term and the assignment to treatment ($cov(z_{i1}, u_{i1})$ in section 3.1.2) can be expected to be negative, which biases the IV estimator downwards.

Based on the exogeneity tests for receiving the BDH and the assignment to treatment, it cannot be assumed that the initial difference is non-existent. In fact, it seems that there are systematic differences between the group of households that receive the BDH and those that do not. Likewise, the households assigned to treatment seem to be worse off than the households assigned to the control group. In a cross sectional analysis the counterfactual for treated households would be the situation of the control households in time 1, y(t=1,C) in Figure 2. Therefore, the OLS and ITT estimators would be just the difference between the two groups in time 1, referred as Diff 1, which is negative for the example. However, given there is an initial difference (Diff 0), the estimator of the effect also includes this systematic difference, which biases the treatment effect negatively in the example.

In summary, the three estimators that use only cross sectional information are not reliable given the evidence of probable endogeneity of both the treatment and the assignment to treatment. Therefore, I suggest using panel data models to overcome the found issues. Their assumptions are not testable, although making use of the three surveys later it can be found traces of reliability on using them. The specifications of the panel models I use rely on the source of bias being additive and time invariant.

6.2 Results of the 1st follow up survey

In Table 2 the results of the different estimation methods are presented. Note that all variables were set such that a positive parameter means an improvement in order to make analysis and interpretation easier to understand. Under regular OLS estimation there are four different estimates that were found to be significant. However, all of them point towards a negative effect of the cash transfer on the situation in the household. Nevertheless, recalling the differences found in the baseline characteristics, these results reflect the treatment effect plus the systematic differences between the two groups, which was in disadvantage to treated households.

Likewise, for the IV and ITT estimation results most of the parameters are related to what was found initially with respect to the systematic differences between both groups. Nevertheless, one result of the IV estimation differs from the OLS findings. A significant increase in the self-reported position in the country is found (MladderE). It is difficult to think of this estimate reflecting intrinsic differences between the groups. However, for this result to hold, it is required the assignment to treatment not to be correlated with other unobserved variables that determine the self-reported position in the country.

The first stage coefficient of the 2SLS estimation is 0.7831 with an F-stat of 916.27[18], which reflects that the estimation issues are not based on a weak instrument problem, but rather on the lack of exogeneity of the instrumental variable. The 2SLS parameters can be estimated as the Wald estimator by dividing the corresponding ITT estimator to the first stage parameter (i.e. 0.783).

Of the three specifications that exploit the panel nature of the data, the first noteworthy point is that in general the parameters of the three specifications are highly similar. The DiD, FE and PSM-DiD estimators take care of systematic

[18] A rule of thumb about strong instruments is that an instrumental variable with an F-stat>10 can be considered as a strong instrument (Staiger & Stock, 1997).

differences that can be present. Once such systemic differences are eliminated, all the significant effects found in the OLS and IV/ITT framework become insignificant and actually no effect was found. It is important to recall that consistency and unbiasedness requires the assumption that the only source of bias comes from the eliminated time invariant unobserved heterogeneity. Some variables are not collected in the baseline, which limits the effects that can be estimated via the DiD, FE and PSM-DiD frameworks.

The probit model for the PSM-DiD model to estimate the propensity score is shown in the Annex 3. The explanation based on the Pseudo-R^2 is 0.100. Almost no covariate is significant, which poses some doubts on the construction of the propensity score, but jointly they are different from zero (Chi2=76.61, p(Chi2)=0,001). However, it is important to remember that the construction of the propensity score is with the objective of comparing more alike observations in their observable characteristics. The first difference $(y_{i1}^{T} - y_{i0}^{T})$ already eliminates the individual fixed effect and the PSM is used to make the comparison group as alike as possible given the observed covariates. The Probit model was obtained based on the baseline characteristics and considering only the observations in the treated parishes. It was used the baseline stated of the outcome variables as well as the control variables in the vector x_i as covariates for the probit model.

Table 2

Estimated effects of BDH

	OLS	IV(2SLS)	ITT	DiD	FE	PSM-DiD
mat	-0.079	-0.120	-0.094	0.002	0.009	-0.035
	(0.067)	(0.102)	(0.081)	(0.061)	(0.06)	(0.051)
floor_d	-0.050	-0.062	-0.048	-0.001	0.002	-0.013
	(0.029)	(0.041)	(0.033)	(0.023)	(0.022)	(0.023)
roof_d	-0.007	-0.013	-0.010	0.017	0.017	0.007
	(0.008)	(0.012)	(0.01)	(0.012)	(0.012)	(0.009)
wall_d	-0.064*	-0.084	-0.066	-0.027	-0.022	-0.036
	(0.03)	(0.046)	(0.037)	(0.021)	(0.02)	(0.021)
matdef	-0.095*	-0.127*	-0.099*	-0.005	0.002	-0.021
	(0.038)	(0.053)	(0.044)	(0.024)	(0.024)	(0.027)
Mhometot	-0.288	-0.388	-0.304	0.248	0.254	0.164
	(0.159)	(0.29)	(0.235)	(0.28)	(0.275)	(0.187)
MladderC	0.276	0.220	0.173			
	(0.162)	(0.263)	(0.211)			
MladderE	0.097	0.347*	0.272			
	(0.116)	(0.173)	(0.141)			
Mworkpay	0.007	0.063	0.050			
	(0.035)	(0.057)	(0.045)			
SAN	-0.063	-0.149	-0.116	0.159	0.188	0.062
	(0.092)	(0.141)	(0.112)	(0.183)	(0.18)	(0.111)
SANexcr_	-0.032	-0.105**	-0.082**	0.001	0.007	-0.010
	(0.024)	(0.036)	(0.028)	(0.052)	(0.051)	(0.036)
SANgarb_	0.020	0.038	0.029	0.079	0.088	0.046
	(0.031)	(0.044)	(0.036)	(0.058)	(0.057)	(0.036)
SANvent_	-0.061*	-0.112**	-0.088*	-0.049	-0.049	-0.022

Table 2 (Continue)

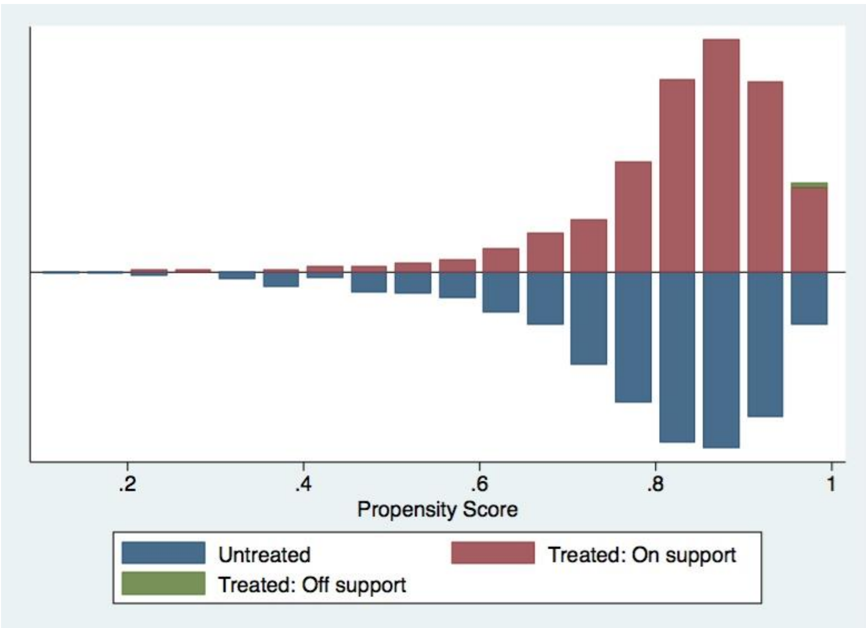
	OLS	IV(2SLS)	ITT	DiD	FE	PSM-DiD
	(0.025)	(0.043)	(0.035)	(0.056)	(0.055)	(0.035)
SANwatr_	-0.018	-0.037	-0.029	0.093	0.097	0.102**
	(0.023)	(0.037)	(0.029)	(0.052)	(0.052)	(0.029)
SANyard_	0.086*	0.138*	0.108*	0.031	0.022	0.074*
	(0.037)	(0.061)	(0.05)	(0.059)	(0.058)	(0.037)
Mpss	-0.340	-0.488	-0.382			
	(0.199)	(0.273)	(0.221)			
Control Vars						
Mom's year of birth 2nd order	x	x	x	x		
Mom's education	x	x	x	x	x	
Household size	x	x	x	x	x	
Mom's head of hh	x	x	x	x	x	
Mom's marital status	x	x	x	x	x	
City fixed effects	x	x	x	x		

**Significant to the 1%, *Significant to the 5%

In parenthesis: clustered standard errors at the parish level, except for the PSM-DiD where 200 replications of bootstrap were used to estimate the standard errors

The common support of treated and control households can be seen in Figure 3. The largest number of observations lies between 0.6 and 0.9, away from 1 and 0 required to successfully run a matching procedure (Abadie & Imbens, 2009). Observations with positive weight are those that are under the common support. There are few treated observations with a large propensity score that are outside the common support.

Figure 3



After the refinement given by the PSM-DiD two effects become significant and go in the direction of the theoretical framework. A significant improvement is found for two of the components of salubrity surrounding the dwelling. It is important to recall that the effects are not the Average Treatment Effect (ATE), but the Average Treatment Effect on the Treated (ATET). Two other matching procedures were also estimated that are found in the Annex 3 as a robustness check for the PSM-DiD framework. The results are sensitive somehow to the choice of matching method. Furthermore, the bootstrapped standard errors are not reliable (Abadie & Imbens, 2008). Therefore, the PSM-DiD estimators are estimated and presented only as an empirical exercise, and the analysis focuses on the DiD and FE results.

6.3 Analysis on cumulative effects

Some policies take longer time to mature, and it is also very likely that the effects of some policies do not only affect a household as a one time shock, but as a continuous shock that keeps improving its conditions.

In the case of a CCT, if there is a permanent increment of the available income in the household, then it can be expected that its effects keep piling up over time. For instance, if a household that has deficient wall and floor decides to use the new income in improving the dwelling characteristics, it is probable that the family won't improve both at the same time but one first and then the other. Therefore, if a survey is collected at any point in time, it is possible to encounter that the household has only improved the floor, but not the wall; however, if a survey is collected some time later it can be found that both the floor and the wall have been improved. Furthermore, each improvement takes time and it is possible that at the time of the survey the improvement has just started, which would be reflected in the survey as no improvement at all. This raises the question of whether two years (the time between the baseline and the 1st follow up survey) is enough time for the effects of the BDH to be observed.

Fortunately, the dataset counts with a second follow up (FU) survey to investigate this point that was collected between May and July 2008. However, for the 2nd FU survey there has been some changes in the structure of the beneficiaries. The design of the randomized experiment was such that the government would not offer the program in the control parishes until the 1st FU survey was collected. After the 1st FU survey, the program also became available in the control parishes, which boosted the take up of the BDH in these parishes. Furthermore, some of the initially treated households stopped receiving the transfer.

Specifically, the dynamics of the transfer are those shown in Table 3. Four different groups were present at the time the second FU was collected:

- Control group (Never received the BDH). 295 households (hh).

- Early treated (Started receiving the BDH before the 1st FU survey). 640 hh.
- Lately treated (Started receiving the BDH after the 1st FU survey). 248 hh.
- Temporarily treated (Started receiving the BDH before the 1st FU survey, but between the 1st and 2nd FU surveys they stopped receiving it) 44 hh.

Table 3

		Receives BDH in 2nd FU survey		
		No	**Yes**	**Total**
Receives BDH in the 1st FU survey	**No**	295	248	543
	Yes	44	640	684
	Total	339	888	1,227

The first specification to capture the effects of the BDH is applying the FE model of equation (15) to the three surveys. This specification contemplates homogenous effects across all the described groups and a constant time trend. This assumption might be hard to sustain in the face of the group of households that stopped receiving the transfer, and the fact that in 2007 the transfer was doubled to USD 30, which might imply heterogeneous effects across groups and across time. In order to capture the probable heterogeneous effects, the following Difference in Differences framework is introduced:

t \ T	Control (C)	Early (E)	Lately (L)	Temporarily (T)	d E/C	d L/C	d T/C
0	γ	α	μ	ρ	$\alpha-\gamma$	$\mu-\gamma$	$\rho-\gamma$
1	$\gamma+\delta$	$\alpha+\delta+\beta_1$	$\mu+\delta+\theta_1$	$\rho+\delta+\varphi_1$	$\beta_1+\alpha-\gamma$	$\theta_1+\mu-\gamma$	$\varphi_1+\rho-\gamma$
2	$\gamma+\lambda$	$\alpha+\lambda+\beta_2$	$\mu+\lambda+\theta_2$	$\rho+\lambda+\varphi_2$	$\beta_2+\alpha-\gamma$	$\theta_2+\mu-\gamma$	$\varphi_2+\rho-\gamma$
d 0/1	δ	$\delta+\beta_1$	$\delta+\theta_1$	$\delta+\varphi_1$	$\boldsymbol{\beta_1}$	$\boldsymbol{\theta_1}$	$\boldsymbol{\varphi_1}$
d 0/2	λ	$\lambda+\beta_2$	$\lambda+\theta_2$	$\lambda+\varphi_2$	$\boldsymbol{\beta_2}$	$\boldsymbol{\theta_2}$	$\boldsymbol{\varphi_2}$

Under this framework, it is allowed to have not only heterogeneous effects based on the treated group, but also based on the time period. β_t, θ_t and φ_t for t=1,2 are the parameters of interest and can be estimated via Difference in Differences using the following specification:

Define the new set of variables

- B_{ei}: Dummy variable if household i is early treated
- B_{li}: Dummy variable if household i is lately treated

- B_{ti}: Dummy variable if household i is temporarily treated
- Y_{it}: outcome of interest for household i in time t
- x_{it}: vector of other K relevant explanatory time invariant variables
- t_{1t}: dummy variable if time is in the 1st FU survey
- t_{2t}: dummy variable if time is in the 2nd FU survey

$$\begin{gathered} Y_{it} = \gamma + \delta * t_{1t} + \lambda * t_{2t} \\ +(\alpha - \gamma) * B_{ei} + \beta_1 * B_{ei} * t_{1t} + \beta_2 * B_{ei} * t_{2t} \\ +(\mu - \gamma) * B_{li} + \theta_1 * B_{li} * t_{1t} + \theta_2 * B_{li} * t_{2t} \\ +(\rho - \gamma) * B_{ti} + \varphi_1 * B_{ti} * t_{1t} + \varphi_2 * B_{ti} * t_{2t} \\ +\tau' * x_{it} + u_{it} \end{gathered} \tag{20}$$

Where, γ, α, μ and ρ are the group fixed effects, δ and λ are the common trends from period 0 to 1 and 1 to 2, respectively. The interpretation of the estimates is for instance β_1 represents the effect of the BDH until the 1st FU survey, β_2 the aggregate effect until the 2nd FU survey, and therefore $\beta_2 - \beta_1$ is the incremental effect from the 1st to the 2nd FU surveys for the early treated group.

In order to gain in consistency if there are individual fixed effects, the same model is estimated via Fixed Effects based on the within estimator of equation (15). Since all the time invariant variables drop out, an OLS estimation is performed in the following specification:

$$\begin{gathered} Y_{it} - \bar{Y}_i = \delta * (t_{1t} - \bar{t_1}) + \lambda * (t_{2t} - \bar{t_2}) \\ +\beta_1 * B_{ei} * (t_{1t} - \bar{t_1}) + \beta_2 * B_{ei} * (t_{2t} - \bar{t_2}) \\ +\theta_1 * B_{li} * (t_{1t} - \bar{t_1}) + \theta_2 * B_{li} * (t_{2t} - \bar{t_2}) \\ +\varphi_1 * B_{ti} * (t_{1t} - \bar{t_1}) + \varphi_2 * B_{ti} * (t_{2t} - \bar{t_2}) \\ +\tau' * (x_{it} - \bar{x}_i) + (u_{it} - \bar{u}_i) \end{gathered} \tag{21}$$

It is important to recall that all time invariant variables drop out including the year of birth of the mother and the municipality fixed effects in the vector x_{it}.

This model not only does it allow to explore heterogeneous group and time effects, but also test the identification assumption of common trend under the DiD framework. If the common trend assumption holds, all the estimates for θ_1 must be zero as it represents the effect of being lately treated until the 1st FU survey. Since they were not treated until after the 1st FU survey, no effect should be found for this group during the first survey. Likewise, considering the temporarily treated group stopped receiving the transfer after the first survey, the estimates of φ_2 should be equal to φ_1. Nevertheless, for the latter to hold, they should have stopped receiving the transfer right after the 1st FU and having received the BDH should not imply a permanent shift in income and consumption that shifts the trend in time of this group.

Based on that, I later estimate a restricted model where I set the restrictions $\theta_1 = 0$ and $\varphi_2 = \varphi_1$.

There are some variables that were not collected during the 1st FU and are only available in the baseline and the 2nd FU survey. For those variables, the effect until the 1st FU survey cannot be estimated and its estimates reflect the aggregate effect from the baseline until the 2nd FU survey only.

The effects were estimated also using the PSM-DiD framework. For this case, the propensity score was calculated estimating the probability of receiving the BDH in the 2nd FU survey based on the baseline characteristics for all the households. Based on that propensity score, the effects for the early and lately treated are estimated up to the 1st and 2nd follow surveys. Given the small number of observations in the temporarily treated group, the standard errors for that group could not be obtained and therefore their results are not presented. Given the arbitrariness when estimating the effects via the PSM-DiD approach, the results are presented in the Annex 5 as a reference only.

6.4 Results of the analysis with the three surveys

Unfortunately, there is no evidence of a net increase in income or negative incentives to adult labour, so the underlying mechanisms explained in Figure 1 of the effects cannot be precisely tracked.

For the heterogeneous model, both the DiD and FE models gave similar results, so for the sake of space I only present the results of the latter, and the DiD results can be found in Annex 4. Furthermore, the FE model is expected to have an advantage over the DiD as it gets rid of more possible sources of bias than the DiD estimator.

The first results are the general Fixed Effects model in column 1 of Table 4, which considers homogenous effects of the BDH across the three different treated groups and a linear time trend. There is a significant improvement of the HOME environment measured by the HOME score. Likewise, a better chance of not finding stagnated water around the house is found.

The questions about assets and access to sanitation were only collected in the baseline and 2nd FU surveys, and therefore their effect can only be measured as an aggregate measure until the 2nd FU survey.

On the assets questions, it was found that the BDH decreased the likelihood of owning a DVD player. This result is not necessarily explained by theory, but can derive from an effect on the reallocation of resources. Treated households might allocate the use of their income in a different way than control households. Recalling the study made by Schady & Rosero (2008), they found households that receive the BDH shift their expenditure towards a larger share in food. This reallocation implies compared to the control households they might be spending less in equipment, or at least some kind of equipment.

When analysing the effects for the different treated groups, interesting differences across the groups and time are found. In general there is an improvement in the home environment measured by the HOME score for two of the three groups: the early and the temporarily treated. Later under the restricted model, the three become significant at least at the 10% level. When analysed via the DiD framework, the three groups already register a positive improvement in the home environment.

Table 4

Estimated effect of BDH
Unrestricted Fixed Effects Model

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	General	Early_1	Early_2	Temp_1	Temp_2	Late_1	Late_2
mat	-0.045	0.000	-0.059	0.047	0.300	-0.018	-0.121
	(0.05)	(0.062)	(0.119)	(0.125)	(0.217)	(0.057)	(0.114)
dwell	-0.044		-0.022		0.251		0.011
	(0.085)		(0.104)		(0.237)		(0.127)
floor_d	0.000	0.011	0.027	0.078	0.080	0.034	0.030
	(0.017)	(0.022)	(0.032)	(0.063)	(0.071)	(0.024)	(0.031)
roof_d	0.007	0.020	0.019	0.015	0.050	0.005	-0.003
	(0.008)	(0.013)	(0.014)	(0.025)	(0.04)	(0.011)	(0.01)
wall_d	-0.021	-0.029	-0.063	0.061	0.179*	-0.002	-0.065
	(0.022)	(0.019)	(0.063)	(0.049)	(0.074)	(0.026)	(0.061)
matdef	0.019	0.002	0.017	0.133+	0.207*	0.023	0.047
	(0.02)	(0.027)	(0.041)	(0.069)	(0.097)	(0.029)	(0.042)
water_d	0.001		-0.011		-0.149+		-0.033
	(0.049)		(0.058)		(0.077)		(0.057)
toilet_d	0.027		0.011		0.089		0.107*
	(0.046)		(0.053)		(0.104)		(0.054)
serdef	0.021		0.003		-0.094		0.028
	(0.05)		(0.06)		(0.113)		(0.057)
hac	-0.013		0.000		-0.074		-0.078
	(0.04)		(0.046)		(0.084)		(0.051)
Mhometot	0.42*	0.112	0.879*	0.890	1.508+	-0.238	0.445
	(0.196)	(0.293)	(0.368)	(0.551)	(0.779)	(0.246)	(0.277)
asset	0.022		0.078		0.126		-0.065
	(0.129)		(0.157)		(0.227)		(0.146)
DRblender	0.046		0.037		-0.011		0.067
	(0.033)		(0.041)		(0.081)		(0.046)
DRcar	-0.019		-0.018		0.013		-0.015
	(0.02)		(0.024)		(0.039)		(0.021)
DRcolortv	0.067		0.095*		-0.041		-0.022
	(0.041)		(0.047)		(0.088)		(0.058)
DRcomputer	-0.011		-0.021		-0.05*		-0.009
	(0.019)		(0.022)		(0.02)		(0.022)
DRfridge	0.012		0.024		-0.008		-0.023
	(0.036)		(0.041)		(0.091)		(0.044)

Table 4 (Continue)

	(1) General	(2) Early_1	(3) Early_2	(4) Temp_1	(5) Temp_2	(6) Late_1	(7) Late_2
DRiron	-0.058		-0.046		0.083		-0.053
	(0.042)		(0.048)		(0.109)		(0.057)
DRmixer	-0.018		-0.032		-0.037		0.003
	(0.026)		(0.03)		(0.067)		(0.034)
DRstereo	-0.023		0.003		0.116		-0.037
	(0.038)		(0.046)		(0.112)		(0.041)
DRstove	0.015		0.012		0.063		0.049*
	(0.017)		(0.021)		(0.05)		(0.022)
DRvhsdvd	-0.116**		-0.104**		-0.066		-0.178**
	(0.037)		(0.039)		(0.111)		(0.049)
DRwasher	-0.038		-0.028		0.025		-0.053*
	(0.025)		(0.033)		(0.052)		(0.024)
SAN	0.140	0.327	0.43+	0.248	0.326	0.317	0.279
	(0.121)	(0.228)	(0.234)	(0.405)	(0.358)	(0.208)	(0.202)
SANexcr_	0.008	0.037	0.105+	0.113	0.130	0.079	0.040
	(0.034)	(0.06)	(0.06)	(0.114)	(0.104)	(0.063)	(0.055)
SANgarb_	0.07+	0.095	0.106	0.112	0.097	0.026	0.009
	(0.038)	(0.072)	(0.068)	(0.126)	(0.13)	(0.065)	(0.063)
SANvent_	0.016	-0.132*	-0.024	0.068	0.059	-0.151*	-0.156*
	(0.042)	(0.06)	(0.066)	(0.117)	(0.13)	(0.065)	(0.07)
SANwatr_	0.099**	0.116+	0.135*	0.022	0.108	0.034	0.060
	(0.036)	(0.062)	(0.065)	(0.127)	(0.131)	(0.047)	(0.049)
SANyard_	-0.035	0.004	-0.092	-0.066	-0.037	-0.058	-0.100
	(0.041)	(0.069)	(0.063)	(0.13)	(0.113)	(0.069)	(0.069)
Control Vars							
Mom's education	x	x	x	x	x	x	x
Household size	x	x	x	x	x	x	x
Mom's head of hh	x	x	x	x	x	x	x
Mom's marital status	x	x	x	x	x	x	x

**Significant to the 1%, *Significant to the 5%, + Significant to the 10%
In parenthesis: Standard errors clustered at the parish level

Table 5

Estimated effect of BDH
Restricted Fixed Effects Model

	(1) Early_1	(2) Early_2	(3) Temp_1	(4) Late_2
mat	0.017	-0.075	0.170	-0.136
	(0.064)	(0.118)	(0.147)	(0.106)
dwell		-0.022	0.251	0.011
		(0.104)	(0.237)	(0.127)
floor_d	-0.004	0.018	0.067	0.013
	(0.023)	(0.032)	(0.062)	(0.029)
roof_d	0.019	0.015	0.030	-0.008
	(0.013)	(0.013)	(0.029)	(0.008)
wall_d	-0.024	-0.071	0.116*	-0.075
	(0.021)	(0.062)	(0.051)	(0.058)
matdef	-0.005	0.005	0.159*	0.028
	(0.026)	(0.041)	(0.073)	(0.04)
water_d		-0.011	-0.149+	-0.033
		(0.058)	(0.077)	(0.057)
toilet_d		0.011	0.089	0.107*
		(0.053)	(0.104)	(0.054)

Table 5 (Continue)

	(1) Early_1	(2) Early_2	(3) Temp_1	(4) Late_2
serdef		0.003	-0.094	0.028
		(0.06)	(0.113)	(0.057)
hac		0.000	-0.074	-0.078
		(0.046)	(0.084)	(0.051)
Mhometot	0.240	0.889**	1.259*	0.51+
	(0.281)	(0.326)	(0.588)	(0.286)
asset		0.078	0.126	-0.065
		(0.157)	(0.227)	(0.146)
DRblender		0.037	-0.011	0.067
		(0.041)	(0.081)	(0.046)
DRcar		-0.018	0.013	-0.015
		(0.024)	(0.039)	(0.021)
DRcolortv		0.095*	-0.041	-0.022
		(0.047)	(0.088)	(0.058)
DRcomputer		-0.021	-0.05*	-0.009
		(0.022)	(0.02)	(0.022)
DRfridge		0.024	-0.008	-0.023
		(0.041)	(0.091)	(0.044)
DRiron		-0.046	0.083	-0.053
		(0.048)	(0.109)	(0.057)
DRmixer		-0.032	-0.037	0.003
		(0.03)	(0.067)	(0.034)
DRstereo		0.003	0.116	-0.037
		(0.046)	(0.112)	(0.041)
DRstove		0.012	0.063	0.049*
		(0.021)	(0.05)	(0.022)
DRvhsdvd		-0.104**	-0.066	-0.178**
		(0.039)	(0.111)	(0.049)
DRwasher		-0.028	0.025	-0.053*
		(0.033)	(0.052)	(0.024)
SAN	0.188	0.347	0.173	0.106
	(0.179)	(0.213)	(0.322)	(0.157)
SANexcr_	0.002	0.085	0.093	-0.003
	(0.052)	(0.055)	(0.097)	(0.038)
SANgarb_	0.083	0.100	0.096	-0.003
	(0.059)	(0.062)	(0.106)	(0.053)
SANvent_	-0.064	0.013	0.117	-0.077
	(0.055)	(0.062)	(0.109)	(0.054)
SANwatr_	0.104*	0.12+	0.050	0.035
	(0.048)	(0.064)	(0.111)	(0.042)
SANyard_	0.031	-0.080	-0.032	-0.073
	(0.056)	(0.058)	(0.103)	(0.059)
Control Vars				
Mom's education	x	x	x	x
Household size	x	x	x	x
Mom's head of hh	x	x	x	x
Mom's marital status	x	x	x	x

**Significant to the 1%, *Significant to the 5%, + Significant to the 10%
All standard errors are clustered at the parish level

The early treated group present the largest effect in the HOME score, but as seen in columns 2 and 3 of Table 4 the effect is cumulative over time. In fact, the effect is not large enough up to the first FU survey, and recently in the 2nd FU survey the effect becomes large enough to be found significant.

Other significant effects for this group are a better salubrity surrounding the house, a higher likelihood of having a Colour TV and less likelihood of possessing a DVD player.

Interestingly, for the temporarily treated households other results are found in columns 4 and 5 of Table 4. It is still found, although only at the 10% significance level, an improvement in the home environment. It is also found that the general conditions of the construction materials, specially driven by the improvement in the condition of the walls, improved for this group. In the sanitation dimension, it is found a worse performance than the control group in the quality of water these households receive. The quality of water is not a variable that is under the control of people. Receiving proper water is a responsibility of public policy, and if anything this shows there has been a better public policy for the control households. It is necessary to point out that this group represents the smallest group of the four. It is only represented by 44 households, which can make the estimates susceptible to extreme values of single observations. Another relevant point is that the effect seems to be cumulative for this group even though they stopped receiving the transfer somewhere between the 1st and the 2nd FU surveys. The dataset does not include the date when they stopped receiving the transfer, so it cannot be tell how long before the 2nd FU survey was collected they did stop receiving the transfer. If they kept receiving the BDH for long time after the 1st FU survey, but stopped receiving it just before the 2nd FU survey, then it would explain the accumulation of effects; nevertheless, the other option is that there was a permanent shift in income that allow this households to keep improving. There is no evidence and no means of testing any of the former hypotheses, which limits the association to the underlying mechanisms. It is important to mention that for this group the larger effects were found and it makes sense with the fact of stop receiving the transfer as they might surpass the cut-point in the SELBEN index.

Furthermore, the lately treated group presents a significant improvement in the adequacy of the toilet in comparison with the control group as seen in the column 7 of Table 4. The use of the group is also the test of the common trend assumption for the method and reassuring the effects are not spurious correlations. Under the common trend assumption, no effect should be found for this group in the first FU survey. As shown by the results in column 6 of Table 4, all variables except of one have insignificant effects. To some extend this helps trusting the control group as a good reference. Since the estimation was done equation by equation, it cannot be tested whether jointly they are zero. However, for the outcomes were more important effects were found, the common trend assumption holds. Such assumption cannot be tested for the other treated groups, as there is not information prior they started receiving the transfer.

The same model was estimated based on the restrictions $\boldsymbol{\theta_1} = \mathbf{0}$ and $\boldsymbol{\varphi_2} = \boldsymbol{\varphi_1}$ in Table 5. In general the same effects as in the unrestricted model were found, although for this case the effect on the home environment for the lately treated becomes significant at the 10% significance level.

Finally, The PSM-DiD results found no effect up to the first FU survey when analysing the two[19] different treatment groups, except for two variables of salubrity under one of the weighting methods for the lately treated. These results show actually how sensitive the PSM-DiD method is to estimate the effects. In the aggregated case, it finds a significant increase in the HOME score for the early treated and no effect for the lately treated group. Nevertheless, it also estimated the treatment worsens the conditions of the toilet and in general of the dwelling of the early treated. In the case of the lately treated it estimates a general negative impact on the possession of assets. The results of this estimation are in the Annex 5.

6.5 Missing observations

So far the analysis has focused on the internal validity of the results; however, besides requiring the unconfoundedness assumption, to correctly identify the treatment effect, the observations with missing values should also be random. Otherwise the external validity of results is in doubt. It can be the case that if observations are not missing at random, the treatment effect on them might be different to the effect on the observations in the used sample, and therefore the results are only valid for the used sample.

A test of means without control variables and a probit model were estimated to compare the baseline characteristics of the groups with and without complete information. According to the first test, both groups are highly different, and systematically the households with missing information are worse off than the sample I use. The probit model finds significant differences in the age squared of the mother and the education of the mother. It cannot be rejected the hypothesis that missing information is not random. It is important, however, to point out that both groups are not different in the relative weight of observations in treated and control parishes. The results can be found in the Annex 2.

Summarizing, given the lack of missing at random, the results cannot be validated externally, and given the systematic differences between treated and non-treated

[19] Because of lack of observations, the effects could not be estimated for the temporarily treated group.

the estimated effects can be interpreted as the average treatment effect on the treated (ATET) rather than an average treatment effect (ATE).

7. Conclusions

In this study I analysed how an unconditional cash transfer in Ecuador called BDH impacts the environment inside the household for a child to correctly develop. Previous studies focused on the effect BDH has on the development of children and rather modest results were found. This study focused on the previous step and checked if the BDH helps building a better environment for children to later be well developed.

The data I use comes from a randomized experiment, but for the sample I use the methods that trust in the randomization of treatment were proven no to be suitable for this study. Therefore, three econometric methods that exploit the panel nature of the data were introduced and are the base of the study: Fixed Effects, Difference in Differences and Propensity Score Matching - Difference in Differences are the methods proposed by this study.

At the end the effects of the transfer are at most modest in improving the general environment of the household. Nevertheless, it was consistently found a significant effect on the measure of the quality of the environment for children measured by the HOME score. I made use of the delay in treatment to some of the households to test the common trend assumption. It was found that for the group that was lately treated the assumption holds and that the found effect can be interpreted as causal. For the other treated groups the assumption cannot be tested and therefore it is an assumption that has to be imposed in the light of its suitability for the lately treated group.

It is relevant to mention that these effects didn't show up in the 1st follow up survey, but they were just visible when addressing the same question using the three surveys at the same time. In fact, it should be expected the effects to be cumulative under the assumption of a permanent shift in the disposable income. This indicates that the effects of a cash transfer might take longer than two years to be visible. Nevertheless, in 2007 there was a change in the amount of the transfer and therefore concluding it was just the pilling up effect can be misleading as this effect can be derived from the increment in the transfer from USD 15 to USD 30.

In general, the effects obtained in this study and some other studies are at most modest, which should guide the country towards rethinking the policy. At this moment (2013), the conditions are different: the amount is USD 50 and the conditionality is expected to be enforced. Given these new conditions, it is recommendable to evaluate if the effectiveness of the policy has improved.

References

Abadie, A., & Imbens, G. W. (2008). On the Failure of the Bootstrap for Matching Estimators. *Econometrica*, *76*(6), 1537–1557. doi:10.3982/ECTA6474

Abadie, A., & Imbens, G. W. (2009). Matching on the Estimated Propensity Score. Cambridge, MA 02138.

Al kalaldeh, M., & Abu Shosha, G. (2012). APPLICATION OF THE PERCEIVED STRESS SCALE IN HEALTH CARE STUDIES . AN ANALYSIS OF LITERATURE. *International Journal of Academic Research*, *4*(4), 45–50.

Araujo, C., & Schady, N. (2008). Cash Transfers, Conditions, and School Enrollment in Ecuador. *Journal of the Latin American and Caribbean Economic Association*, *8*(2), 43–77.

Bradley, R. H., Caldwell, B. M., Rock, S. L., Gottfried, A. W., Barnard, K. E., Gray, C., … Johnson, D. L. (1989). Home Environment and Cognitive Development in the First 3 Years of Life: A Collaborative Study Involving Six Sites and Three Ethnic Groups in North America. *Developmentat Psychology*, *25*(2), 217–235.

Caldwell, B. M., & Bradley, R. H. (2003). Home Observation for Measurement of the Environment: Administration Manual. Tempe, AZ: Family & Human Dynamics Research Institute, Arizona State University.

Cameron, A. C., & Trivedi, P. K. (2005). *Microeconometrics: Methods and Applications*. *Analysis* (Vol. 100, p. 1056). Cambridge University Press. doi:10.1016/S0304-4076(00)00050-6

Cohen, S. (1983). A Global Measure of Perceived Stress. *Journal of Health and Social Behavior*, *24*(4), 385–396.

Dehejia, R. H., & Wahba, S. (1998). Propensity score-matching methods for nonexperimental causal studies.

Edmonds, E., & Schady, N. (2012). Poverty Alleviation and Child Labor. *American Economic Journal: Economic Policy*, *4*(4), 100–124.

Felfe, C., & Hsin, A. (2012). Maternal work conditions and child development. *Economics of Education Review*, *31*(6), 1037–1057. doi:10.1016/j.econedurev.2012.06.008

Fernald, L. C. H., & Hidrobo, M. (2011). Effect of Ecuador's cash transfer program (Bono de Desarrollo Humano) on child development in infants and toddlers: a randomized effectiveness trial. *Social science & medicine (1982)*, *72*(9), 1437–46. doi:10.1016/j.socscimed.2011.03.005

Ferron, J., Ng'Andu, N., & Garrett, P. (1994). Evaluating the Dimensional Structure of the Home Observation for Measurement of the Environment-Short form.

Educational and Psychological Measurement, *54*(2), 537–540. doi:10.1177/0013164494054002028

Fiszbein, A., & Schady, N. (2009). *Conditional Cash Transfers*.

Giatti, L., Camelo, L. D. V., Rodrigues, J. F. D. C., & Barreto, S. M. (2012, January). Reliability of the MacArthur scale of subjective social status - Brazilian Longitudinal Study of Adult Health (ELSA-Brasil). *BMC public health*. BMC Public Health. doi:10.1186/1471-2458-12-1096

Grilli, L., & Rampichini, C. (2011). *Presentation on Propensity scores for the estimation of average treatment effects in observational studies*. Bristol.

Handa, S., & Davis, B. (2006). The Experience of Conditional Cash Transfers in Latin America and the Caribbean. *Development Policy Review*, *24*(5), 513–536. doi:10.1111/j.1467-7679.2006.00345.x

Hirano, K., Imbens, G. W., & Ridder, G. (2003). EFFICIENT ESTIMATION OF AVERAGE TREATMENT EFFECTS USING THE ESTIMATED PROPENSITY SCORE. *Econometrica*, *71*(4), 1161–1189. doi:10.1111/1468-0262.00442

Imbens, G. W., & Angrist, J. D. (1994). Identification and Estimation of Local Average Treatment Effects. *Econometrica*, *62*(2), 467–475. doi:10.2307/2951620

Johnson, D. L., Swank, P., Howie, V. M., Baldwin, C. D., Owen, M., & Luttman, D. (1993). Does HOME Add to the Prediction of Child Intelligence Over and Above SES ? *The Journal of Genetic Psychology: Research and Theory on Human Development*, *154*(1), 37–40.

Khandker, S., B. Koolwal, G., & Samad, H. (2009). *Handbook on Impact Evaluation*. World Bank. doi:10.1596/978-0-8213-8028-4

Leuven, E. & Sianesi, B. (2003). "PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing". http://ideas.repec.org/c/boc/bocode/s432001.html.

Leventhal, T., & Newman, S. (2010). Housing and child development. *Children and Youth Services Review*, *32*(9), 1165–1174. doi:10.1016/j.childyouth.2010.03.008

Maluccio, J. A., & Flores, R. (2004). IMPACT EVALUATION OF A CONDITIONAL CASH TRANSFER PROGRAM : THE NICARAGUAN RED DE PROTECCIÓN SOCIAL. Washington D.C., USA.

Nabernegg, M. (2012). *The impact of the Bono de Desarrollo Humano in the expenditure for undesirable goods: A regression discontinuity analysis* (pp. 0–11). Retrieved from http://mpra.ub.uni-muenchen.de/41295/

Paxson, C., & Schady, N. (2005). Cognitive Development Among Young Children in Ecuador: The Roles of Wealth, Health and Parenting. *Cognitive Development*, *42*(1), 49. Retrieved from http://jhr.uwpress.org/cgi/content/abstract/XLII/1/49

Paxson, C., & Schady, N. (2010). Does money matter? The effects of cash transfers on child development in rural Ecuador. *Economic development and cultural change*, *59*(1), 187–229. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/20821896

Ponce, J., & Bedi, A. S. (2010). The impact of a cash transfer program on cognitive achievement: The Bono de Desarrollo Humano of Ecuador. *Economics of Education Review*, *29*(1), 116–125. doi:10.1016/j.econedurev.2009.07.005

Rubin, D. (1974). Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies. *Journal of Educational Psychology*, *66*, 688–701.

Schady, N. (2012). Cash transfers and anemia among women of reproductive age. *Economics Letters*, *117*(3), 887–890. doi:10.1016/j.econlet.2012.07.014

Schady, N., & Rosero, J. (2008). Are cash transfers made to women spent like other sources of income? *Economics Letters*, *101*(3), 246–248. doi:10.1016/j.econlet.2008.08.015

Skoufias, E., & Parker, S. W. (2001). CONDITIONAL CASH TRANSFERS AND THEIR IMPACT ON CHILD WORK AND SCHOOLING: EVIDENCE FROM THE PROGRESA PROGRAM IN MEXICO. Washington D.C., USA.

Staiger, D., & Stock, J. H. (1997). Instrumental Variables Regression with Weak Instruments. *Econometrica*, *65*(3), 557–586. doi:10.2307/2171753

Stock, J., & Watson, M. (2008). *Introduction to econometrics*. (Pearson/Addison-Wesley, Ed.) (3rd ed.). Pearson/Addison-Wesley. Retrieved from http://tocs.ulb.tu-darmstadt.de/98136275.pdf

Totsika, V., & Sylva, K. (2004). The Home Observation for Measurement of the Environment Revisited. *Child and Adolescent Mental Health*, *9*(1), 25–35. doi:10.1046/j.1475-357X.2003.00073.x

Villatoro, P. (2005, August). Revista de la CEPAL 86, 87–101.

Waldfogel, J. (2007). Child Work Arrangements and Child Development. *Canadian Public Policy*, *33*(June 2007), 251–271.

Younger, S. D., Ponce, J., & Hidalgo, D. (2009). El Impacto de Programas de Transferencias a las Madres de Familia en la Seguridad Alimentaria de los Niños : Un análisis comparado de los casos de México y Ecuador.

Annex 1
Construction of Dependent Variables

floor_d: Adequate floor material is: finished wood, tile, marble, cement/brick or floorboard. If the household has floor of cane, dirt or other not enlisted material it is considered inadequate.

roof_d: Adequate roof material is: concrete, asbestos, zinc and tile. Inadequate materials are palm/straw and other not enlisted.

wall_d: Adequate walls material is: concrete/brick and block. Inadequate materials are asbestos, sundried brick, wood and mud-covered cane.

mat: The first factor of a Principal Components Analysis (PCA) between floor_d, roof_d and wall_d. The first component explains 41% of the variance of the three variables. The obtained weights of the first eigenvalue were -0.3928, 0.5967 and 0.6997, respectively.

matdef: The general adequacy of construction materials was calculated as a dummy variable, where 1 indicates if the floor, roof and walls are all adequate. If at least one of them is not adequate, then matdef takes the value of 0.

hac: not in overcrowding. A household is considered not to live in overcrowding conditions if there are less than 3 people per bedroom.

water_d: adequate water source is considered only public network. If the water is obtained from other source then it is considered inadequate.

toilet_d: it is considered adequate toilet if the households has toilet and either is connected to the public network or to a septic tank. Any other form of toilet is considered inadequate.

dwell: it is the first factor of a PCA analysis among floor_d, roof_d, wall_d, hac, water_d and toilet_d. The first factor captures 28% of the variance of the variables. The weights from the first factor's eigenvalue are 0.2339, 0.2892, 0.4731, 0.1503, 0.5221 and 0.5853, respectively.

DRfrige, DRstove, DRblender, DRmixer, DRiron, DRcolortv, DRstereo, DRvhsdvd, DRcar, DRcomputer, DRwasher: are dummy variables that indicate possession of the described asset.

asset: it is the first factor of a PCA analysis among DRfrige, DRstove, DRblender, DRmixer, DRiron, DRcolortv, DRstereo, DRvhsdvd, DRcar, DRcomputer, DRwasher. The first factor captures 25% of the variance across all variables. The weights from the first factor are 0.4021, 0.2342, 0.3877, 0.2540, 0.3594, 0.3954, 0.3332, 0.2629, 0.2466, 0.1667 and 0.1196, respectively.

SANexcr_, SANgarb_, SANwatr_, SANvent_ and SANyard_: are dummy variables indicating if the conditions explained in Table 1 are met.

SAN: it's the first factor of a PCA analysis of SANexcr_, SANgarb_, SANwatr_, SANvent_ and SANyard_. 52% of the variance is captured by the first factor. The weights are 0.4834, 0.5001, -0.3514, 0.3818 and -0.4970, respectively.

All PCA analyses were conducted in the baseline and using the first eigenvector (weights) the factors were constructed for the 1st and 2nd follow up surveys. In order to construct the factors in the 1st and 2nd follow up surveys, the variables need to be standardized.

HOME score: Is a scale of eleven questions concerning parents' responsivity and punitiveness towards their children. The questions were filled by observation and required at least one child to be awake during the interview. The following translation of the questions was obtained from Paxson & Schady (2005).

Responsivity

1. Did the mother or father spontaneously say kind words or phrases to the children at least twice during the interview?
2. At least once, did the mother or father respond verbally to a child's vocalization?
3. At least once, did the mother or father tell the child the name of an object?
4. At least twice, did the mother or father praise one of the children?
5. Did the mother or father convey positive feelings toward the children when they speak to or about them?
6. Did the mother or father caress or kiss one of the children at least once?

Punitiveness

1. Did the mother or father yell at any of the children?
2. Was the mother or father annoyed with or hostile toward any of the children?
3. During the interview, did the mother or father hit any of the children?
4. During the interview, did the mother or father scold or criticize any of the children?
5. Did the mother or father forbid any of the children from doing something more than three times during the interview?

Each question was answered with a yes or a no. If a no was the answer for a question about responsivity, then 1 was added and otherwise 0. If a yes was the answer for a question about punitiveness then 1 was added and otherwise 0. Therefore the scale ranges from 0 to 11. In this thesis, it is used the negative HOME score in order to relate a positive coefficient with an improvement.

MladderC/E: MacArthur ladders are instruments given to the mother, who is asked to point out in a scale from 1 to 9, where 9 is the highest place, where she thinks her household is with respect to her community (C) or in Ecuador (E).

PSS score: it is a perceived stress scale of free use constructed originally by Sheldon Cohen that was adapted to the survey that is used in this thesis and collects four questions of the ten questions of the original scale[20]:

1. In the last month, how often have you felt that you were unable to control the important things in your life?
2. In the last month, how often have you felt confident about your ability to handle your personal problems?
3. In the last month, how often have you felt that things were going your way?
4. In the last month, how often have you felt difficulties were piling up so high that you could not overcome them?

Each question is answered as 0 = Never, 1 = Almost Never, 2 = Sometimes, 3 = Fairly Often or 4 = Very Often.

Then questions 2 and 3 are recoded in a reverse order (0=4) (1=3) (2=2) (3=1) (4=0). Then the scores are summed up to obtain the scale.

[20] The author of the scale himself recommends this selection.

Annex 2
Exogeneity Tests

	Assignment to treatment			BDH			Missing	
	ttest No Controls	ttest with Controls	Probit	ttest No Controls	ttest with Controls	Probit	ttest with Controls	Probit
floor_d0			*		*		***	
roof_d0	*			*	**			
wall_d0							**	
matdef0		*	**	*	**		***	
water_d0							**	
toilet_d0					**		***	
serdef0		*			**		***	
hac0			*	*	**	**		
DRbike0		***					*	
DRblender0							***	
DRcar0		**		**	***	**	**	
DRcolortv0		*		*	**		***	
DRcomputer0	**	*	**	*				
DRfan0								
DRfridge0				**	**	**	***	
DRiron0			***			*	***	
DRmixer0							***	*
DRmotorc0			**					
DRoven0						*		
DRradio0								
DRsewing0								
DRstereo0	**	**	**	**			***	
DRstove0							**	
DRtypewr0		**	**	*	**	**	***	
DRvhsdvd0		**		***	***	***		
DRwaffle0								
DRwasher0							**	
Mhometot0								
SANexcr_0		**			*		*	
SANgarb_0	*	*		**	*			
SANvent_0			**		*			
SANwatr_0	***	***	***	**	*			**
SANyard_0					*		*	
n_0t14_0								
n_15t64_0								*
n_65_0								
Mage0						*	**	
Mage02						**	**	*
Medyrs0			*			***	***	**
treat								

Significance based on Clustered St. Errors. Clustered at Parish Level
*** p<0.01, ** p<0.05, * p<0.10

Annex 3
PSM Estimation for the 1st Follow Up Survey

	Probit bdh1 = f(x)
VARIABLES	
n_0t140	-0.0251
	(0.0583)
n_15t640	-0.00302
	(0.0513)
n_650	0.197
	(0.162)
floor_d0	-0.641
	(0.479)
roof_d0	-0.377
	(0.341)
wall_d0	-0.697
	(0.467)
matdef0	0.535
	(0.485)
water_d0	-0.152
	(0.162)
toilet_d0	-0.101
	(0.186)
serdef0	-0.130
	(0.243)
hac0	-0.136
	(0.131)
DRbike0	0.244
	(0.159)
DRblender0	0.0798
	(0.134)
DRbwtv0	-0.0450
	(0.145)
DRcar0	-1.136***
	(0.423)
DRcolortv0	-0.143
	(0.152)
DRfan0	0.268
	(0.247)
DRfridge0	-0.208
	(0.143)
DRiron0	-0.0148
	(0.135)
DRmixer0	0.0558
	(0.272)
DRmotorc0	-0.0426
	(0.659)
DRoven0	0.895
	(0.574)
DRradio0	0.0563
	(0.132)
DRsewing0	0.0692
	(0.187)
DRstereo0	0.104
	(0.159)
DRstove0	0.291
	(0.209)

Probit bdh1 = f(x)
(Continue)

VARIABLES	
DRtypewr0	-0.534*
	(0.282)
DRvhsdvd0	-0.609***
	(0.207)
Mhometot0	-0.0137
	(0.0261)
SANexcr_0	0.0887
	(0.150)
SANgarb_0	-0.176
	(0.153)
SANvent_0	0.0152
	(0.134)
SANwatr_0	0.279**
	(0.141)
SANyard_0	0.258
	(0.161)
Mage0	-0.0307
	(0.0606)
Mage02	0.000802
	(0.00102)
Medyrs0	0.167**
	(0.0742)
Medyrs02	-0.00677
	(0.00459)
Mlivehusb0	0.270
	(0.258)
Msta02	0.242
	(0.148)
Msta03	0.522
	(0.336)
Msta04	0.544*
	(0.286)
Mhead0	-0.388
	(0.247)
Constant	0.904
	(1.139)
Pseudo-R2	**0.1009**
LR chi2	**76.61**
p(chi2)	**0.0012**
Observations	**806**

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Estimated effects of BDH
Under Different Matching Methods

	LLR+	N5++
mat	-0.041	-0.031
	(0.053)	(0.059)
floor_d	-0.013	-0.005
	(0.023)	(0.026)
roof_d	0.009	0.010
	(0.01)	(0.011)
wall_d	-0.044*	-0.035
	(0.022)	(0.025)
matdef	-0.019	-0.013
	(0.028)	(0.033)
Mhometot	0.213	0.232
	(0.189)	(0.21)
SAN	0.044	0.054
	(0.111)	(0.129)
SANexcr_	-0.018	-0.006
	(0.038)	(0.043)
SANgarb_	0.043	0.040
	(0.037)	(0.042)
SANvent_	-0.023	-0.013
	(0.038)	(0.044)
SANwatr_	0.095**	0.109**
	(0.031)	(0.035)
SANyard_	0.077*	0.081
	(0.038)	(0.043)

**Significant to the 1%, *Significant to the 5%
In parenthesis: Bootstrapped St. Errors with 200 replications
+ LLR: Local Linear Regression
++ N5: 5 closest neighbours

Annex 4
Difference in Differences Estimator for the 2nd Follow Up Survey

Estimated effect of BDH
Unrestricted DiD Model

	(2) Early_1	(3) Early_2	(4) Temp_1	(5) Temp_2	(6) Late_1	(7) Late_2
mat	-0.005	-0.069	0.018	0.235	-0.008	-0.113
	(0.063)	(0.121)	(0.123)	(0.214)	(0.058)	(0.116)
dwell		-0.047		0.349		0.009
		(0.108)		(0.231)		(0.131)
floor_d	0.011	0.025	0.073	0.114	0.034	0.028
	(0.023)	(0.031)	(0.063)	(0.069)	(0.024)	(0.031)
roof_d	0.019	0.019	0.014	0.051	0.005	-0.002
	(0.013)	(0.014)	(0.025)	(0.039)	(0.011)	(0.01)
wall_d	-0.031	-0.069	0.043	0.162*	0.003	-0.063
	(0.02)	(0.063)	(0.048)	(0.074)	(0.027)	(0.062)
matdef	0.000	0.009	0.115	0.223*	0.027	0.046
	(0.028)	(0.041)	(0.071)	(0.099)	(0.03)	(0.043)
water_d		-0.012		-0.101		-0.034
		(0.059)		(0.082)		(0.056)
toilet_d		-0.001		0.135		0.105+
		(0.054)		(0.105)		(0.056)
serdef		-0.007		-0.060		0.025
		(0.061)		(0.114)		(0.058)
hac		-0.011		-0.030		-0.079
		(0.046)		(0.084)		(0.052)
Mhometot	0.110	0.881*	0.878	1.623*	-0.221	0.462+
	(0.294)	(0.371)	(0.542)	(0.802)	(0.244)	(0.28)
asset		0.064		0.372		-0.052
		(0.159)		(0.251)		(0.145)
DRblender		0.040		0.055		0.073
		(0.041)		(0.083)		(0.047)
DRcar		-0.020		0.042		-0.018
		(0.024)		(0.047)		(0.02)
DRcolortv		0.086+		-0.018		-0.019
		(0.048)		(0.09)		(0.059)
DRcomputer		-0.020		-0.034+		-0.012
		(0.023)		(0.02)		(0.022)
DRfridge		0.018		0.033		-0.020
		(0.042)		(0.092)		(0.045)
DRiron		-0.048		0.138		-0.048
		(0.049)		(0.109)		(0.058)
DRmixer		-0.033		-0.014		0.002
		(0.031)		(0.069)		(0.035)
DRstereo		0.002		0.152		-0.036
		(0.046)		(0.11)		(0.041)
DRstove		0.010		0.065		0.052*
		(0.021)		(0.046)		(0.022)
DRvhsdvd		-0.101*		-0.071		-0.176**
		(0.039)		(0.112)		(0.049)
DRwasher		-0.027		0.013		-0.052*
		(0.033)		(0.055)		(0.024)
SAN	0.319	0.417+	0.202	0.332	0.334	0.305
	(0.232)	(0.238)	(0.404)	(0.348)	(0.208)	(0.206)

Estimated effect of BDH
Unrestricted DiD Model (Continue)

	(2) Early_1	(3) Early_2	(4) Temp_1	(5) Temp_2	(6) Late_1	(7) Late_2
SANexcr_	0.036	0.105+	0.099	0.140	0.083	0.046
	(0.061)	(0.06)	(0.115)	(0.104)	(0.063)	(0.056)
SANgarb_	0.092	0.102	0.097	0.112	0.031	0.015
	(0.073)	(0.068)	(0.125)	(0.128)	(0.066)	(0.064)
SANvent_	-0.132*	-0.022	0.059	0.075	-0.152*	-0.161*
	(0.061)	(0.067)	(0.115)	(0.127)	(0.064)	(0.07)
SANwatr_	0.115+	0.134*	0.016	0.097	0.037	0.067
	(0.062)	(0.066)	(0.127)	(0.126)	(0.048)	(0.048)
SANyard_	0.007	-0.086	-0.049	-0.048	-0.062	-0.103
	(0.071)	(0.064)	(0.128)	(0.112)	(0.069)	(0.071)
Control Vars						
Mom's education	x	x	x	x	x	x
Household size	x	x	x	x	x	x
Mom's head of hh	x	x	x	x	x	x
Mom's marital status	x	x	x	x	x	x

**Significant to the 1%, *Significant to the 5%, + Significant to the 10%
In parenthesis: Standard errors clustered at the parish level

Estimated effect of BDH
Restricted DiD Model

	(1) Early_1	(2) Early_2	(3) Temp_1	(4) Late_2
mat	0.007	-0.084	0.123	-0.130
	(0.064)	(0.118)	(0.144)	(0.106)
dwell		-0.047	0.349	0.009
		(0.108)	(0.231)	(0.131)
floor_d	-0.003	0.013	0.080	0.006
	(0.023)	(0.032)	(0.061)	(0.029)
roof_d	0.018	0.015	0.029	-0.009
	(0.013)	(0.013)	(0.028)	(0.008)
wall_d	-0.028	-0.079	0.098*	-0.076
	(0.022)	(0.062)	(0.05)	(0.058)
matdef	-0.009	-0.006	0.157*	0.022
	(0.026)	(0.041)	(0.075)	(0.039)
water_d		-0.012	-0.101	-0.034
		(0.059)	(0.082)	(0.056)
toilet_d		-0.001	0.135	0.105+
		(0.054)	(0.105)	(0.056)
serdef		-0.007	-0.060	0.025
		(0.061)	(0.114)	(0.058)
hac		-0.011	-0.030	-0.079
		(0.046)	(0.084)	(0.052)
Mhometot	0.235	0.876**	1.306*	0.504+
	(0.281)	(0.329)	(0.596)	(0.291)
asset		0.064	0.372	-0.052
		(0.159)	(0.251)	(0.145)
DRblender		0.040	0.055	0.073
		(0.041)	(0.083)	(0.047)

Estimated effect of BDH
Restricted DiD Model (Continue)

	(1) Early_1	(2) Early_2	(3) Temp_1	(4) Late_2
DRcar		-0.020	0.042	-0.018
		(0.024)	(0.047)	(0.02)
DRcolortv		0.086+	-0.018	-0.019
		(0.048)	(0.09)	(0.059)
DRcomputer		-0.020	-0.034+	-0.012
		(0.023)	(0.02)	(0.022)
DRfridge		0.018	0.033	-0.020
		(0.042)	(0.092)	(0.045)
DRiron		-0.048	0.138	-0.048
		(0.049)	(0.109)	(0.058)
DRmixer		-0.033	-0.014	0.002
		(0.031)	(0.069)	(0.035)
DRstereo		0.002	0.152	-0.036
		(0.046)	(0.11)	(0.041)
DRstove		0.010	0.065	0.052*
		(0.021)	(0.046)	(0.022)
DRvhsdvd		-0.101*	-0.071	-0.176**
		(0.039)	(0.112)	(0.049)
DRwasher		-0.027	0.013	-0.052*
		(0.033)	(0.055)	(0.024)
SAN	0.174	0.324	0.147	0.118
	(0.182)	(0.217)	(0.316)	(0.158)
SANexcr_	0.000	0.081	0.089	-0.001
	(0.052)	(0.055)	(0.097)	(0.038)
SANgarb_	0.079	0.093	0.093	-0.003
	(0.06)	(0.063)	(0.104)	(0.052)
SANvent_	-0.063	0.015	0.120	-0.083
	(0.056)	(0.063)	(0.106)	(0.053)
SANwatr_	0.101*	0.118+	0.041	0.039
	(0.049)	(0.065)	(0.11)	(0.042)
SANyard_	0.035	-0.071	-0.026	-0.071
	(0.057)	(0.059)	(0.1)	(0.06)
Control Vars				
Mom's education	x	x	x	x
Household size	x	x	x	x
Mom's head of hh	x	x	x	x
Mom's marital status	x	x	x	x

**Significant to the 1%, *Significant to the 5%, + Significant to the 10%
All standard errors are clustered at the parish level

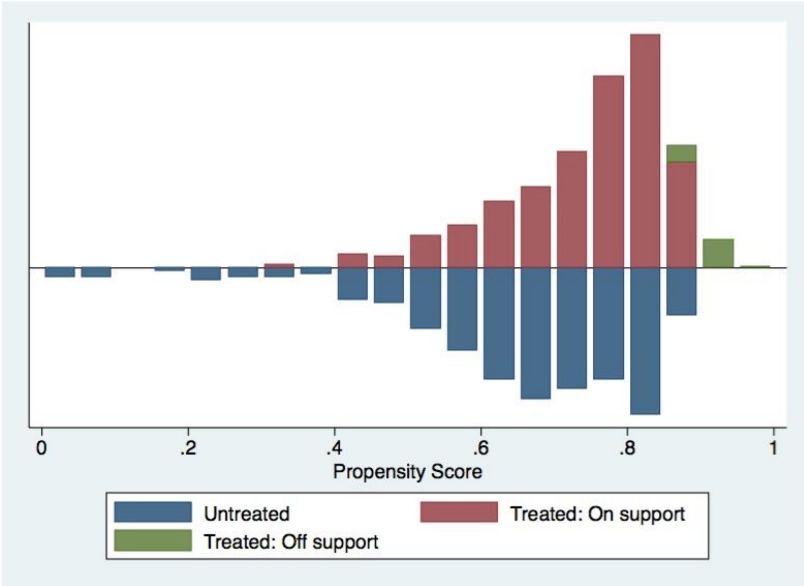
Annex 5
PSM-DiD Estimator for the 2nd Follow Up Survey

	Probit bdh2 = f(x)
VARIABLES	
n_0t140	-0.0146
	(0.0430)
n_15t640	-0.00665
	(0.0385)
n_650	0.127
	(0.110)
floor_d0	0.115
	(0.243)
roof_d0	0.0477
	(0.233)
wall_d0	0.363
	(0.237)
matdef0	-0.289
	(0.261)
water_d0	-0.0734
	(0.117)
toilet_d0	-0.0772
	(0.136)
serdef0	-0.271
	(0.176)
hac0	-0.136
	(0.0952)
DRbike0	0.0150
	(0.112)
DRblender0	-0.0395
	(0.0952)
DRbwtv0	0.0895
	(0.105)
DRcar0	-0.601**
	(0.278)
DRcolortv0	-0.0755
	(0.110)
DRcomputer0	-1.368*
	(0.814)
DRfan0	0.0984
	(0.179)
DRfridge0	-0.202**
	(0.101)
DRiron0	0.209**
	(0.0964)
DRmixer0	-0.0959
	(0.187)
DRmotorc0	-0.617
	(0.569)
DRoven0	-0.0307
	(0.310)
DRradio0	-0.123
	(0.0929)
DRsewing0	0.274**
	(0.136)
DRstereo0	-0.175
	(0.111)

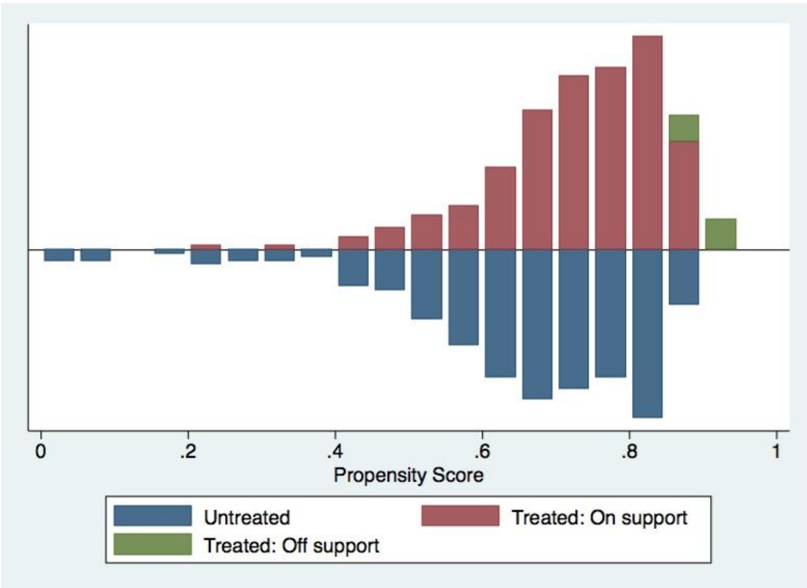
Probit bdh2 = f(x) (Continue)	
VARIABLES	
DRstove0	0.113
	(0.151)
DRtypewr0	-0.169
	(0.197)
DRvhsdvd0	-0.233
	(0.153)
DRwaffle0	0.602*
	(0.358)
DRwasher0	-0.423
	(0.961)
Mhometot0	-0.00918
	(0.0186)
SANexcr_0	-0.0148
	(0.105)
SANgarb_0	-0.0587
	(0.109)
SANvent_0	0.0167
	(0.0922)
SANwatr_0	-0.0329
	(0.103)
SANyard_0	-0.0284
	(0.110)
Mage0	-0.0249
	(0.0373)
Mage02	0.000368
	(0.000595)
Medyrs0	0.0498
	(0.0562)
Medyrs02	-0.00346
	(0.00350)
Mlivehusb0	0.190
	(0.186)
Msta02	-0.212**
	(0.106)
Msta03	-0.0294
	(0.235)
Msta04	0.0675
	(0.201)
Mhead0	-0.0398
	(0.185)
Constant	0.829
	(0.734)
Pseudo-R2	**0.074**
LR chi2	**107.08**
p(chi2)	**0.00**
Observations	**1,227**

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Common Support for early treated



Common Support for lately treated



Estimated effect of BDH

PSM-DiD estimation based on Treatment Group up to the 1st Follow Up Survey

	PSM-DiD_e_k	PSM-DiD_e_l	PSM-DiD_e_n	PSM-DiD_l_k	PSM-DiD_l_l	PSM-DiD_l_n
mat	0.026	0.031	0.030	-0.007	0.008	0.007
	(0.075)	(0.083)	(0.084)	(0.074)	(0.081)	(0.087)
floor_d	-0.009	-0.019	-0.023	0.011	0.008	-0.003
	(0.028)	(0.03)	(0.032)	(0.033)	(0.036)	(0.038)
roof_d	0.017	0.016	0.013	-0.003	0.001	0.001
	(0.017)	(0.019)	(0.019)	(0.013)	(0.015)	(0.016)
wall_d	-0.021	-0.021	-0.018	0.004	0.006	-0.001
	(0.026)	(0.027)	(0.029)	(0.032)	(0.034)	(0.036)
matdef	-0.017	-0.028	-0.029	0.003	-0.001	-0.008
	(0.035)	(0.037)	(0.042)	(0.043)	(0.046)	(0.051)
Mhometot	-0.063	-0.118	-0.155	-0.337	-0.361	-0.290
	(0.201)	(0.233)	(0.263)	(0.231)	(0.251)	(0.302)

Estimated effect of BDH
PSM-DiD estimation based on Treatment Group up to the 1st Follow Up Survey
(Continue)

	PSM-DiD_e_k	PSM-DiD_e_l	PSM-DiD_e_n	PSM-DiD_l_k	PSM-DiD_l_l	PSM-DiD_l_n
SAN	0.143	0.189	0.157	0.193	0.311	0.402*
	(0.129)	(0.155)	(0.173)	(0.163)	(0.174)	(0.191)
SANexcr_	0.029	0.046	0.032	0.081	0.117*	0.136*
	(0.045)	(0.051)	(0.057)	(0.055)	(0.059)	(0.065)
SANgarb_	0.061	0.057	0.056	-0.008	0.003	0.031
	(0.045)	(0.05)	(0.056)	(0.054)	(0.06)	(0.063)
SANvent_	-0.046	-0.051	-0.057	-0.076	-0.094	-0.11*
	(0.042)	(0.051)	(0.051)	(0.047)	(0.051)	(0.055)
SANwatr_	0.057	0.066	0.057	0.003	0.026	0.048
	(0.035)	(0.043)	(0.047)	(0.038)	(0.042)	(0.05)
SANyard_	0.046	0.026	0.036	-0.034	-0.065	-0.077
	(0.042)	(0.05)	(0.055)	(0.054)	(0.056)	(0.062)

**Significant to the 1%, *Significant to the 5%
In parenthesis: Bootstrapped St. Errors with 200 replications

Estimated effect of BDH
Aggregated Effect via PSM-DiD estimation based on Treatment Group up to the 2nd Follow Up Survey

	(2)	(3)	(4)	(5)	(6)	(7)
	PSM-DiD_e_k	PSM-DiD_e_l	PSM-DiD_e_n	PSM-DiD_l_k	PSM-DiD_l_l	PSM-DiD_l_n
mat	-0.001	0.029	0.014	-0.048	0.003	-0.027
	(0.086)	(0.097)	(0.1)	(0.095)	(0.102)	(0.106)
dwell	-0.174+	-0.184+	-0.22+	-0.078	-0.034	-0.074
	(0.104)	(0.112)	(0.118)	(0.119)	(0.12)	(0.134)
floor_d	-0.003	-0.012	-0.009	-0.003	-0.013	-0.003
	(0.031)	(0.032)	(0.036)	(0.038)	(0.041)	(0.044)
roof_d	0.010	0.010	0.010	-0.011	-0.004	-0.004
	(0.013)	(0.017)	(0.017)	(0.009)	(0.009)	(0.01)
wall_d	-0.016	0.000	-0.011	-0.015	0.000	-0.019
	(0.043)	(0.047)	(0.05)	(0.055)	(0.057)	(0.059)
matdef	-0.001	0.001	-0.007	0.030	0.037	0.035
	(0.041)	(0.045)	(0.049)	(0.049)	(0.053)	(0.056)
water_d	-0.052	-0.043	-0.061	-0.061	-0.053	-0.048
	(0.046)	(0.048)	(0.052)	(0.05)	(0.051)	(0.058)
toilet_d	-0.095*	-0.119**	-0.13**	0.042	0.054	0.026
	(0.039)	(0.043)	(0.045)	(0.052)	(0.056)	(0.062)
serdef	-0.060	-0.069	-0.078+	-0.016	0.002	-0.002
	(0.04)	(0.043)	(0.047)	(0.047)	(0.048)	(0.053)
hac	0.012	0.000	0.009	-0.082	-0.073	-0.083
	(0.046)	(0.049)	(0.054)	(0.056)	(0.058)	(0.063)
Mhometot	0.61*	0.603*	0.555+	0.210	0.249	0.386
	(0.243)	(0.267)	(0.302)	(0.287)	(0.303)	(0.344)
asset	-0.226	-0.200	-0.212	-0.313+	-0.345+	-0.416*
	(0.138)	(0.146)	(0.156)	(0.167)	(0.176)	(0.189)
DRblender	-0.005	0.011	-0.001	0.039	0.022	0.013
	(0.048)	(0.054)	(0.057)	(0.051)	(0.057)	(0.058)
DRcar	-0.036	-0.042+	-0.047+	-0.031	-0.038	-0.047+
	(0.023)	(0.025)	(0.027)	(0.024)	(0.024)	(0.026)
DRcolortv	-0.015	-0.012	-0.016	-0.088+	-0.086	-0.093
	(0.039)	(0.041)	(0.042)	(0.051)	(0.054)	(0.06)

Estimated effect of BDH
Aggregated Effect via PSM-DiD estimation based on Treatment Group up to the 2nd Follow Up Survey (Continue)

	(2) PSM-DiD_e_k	(3) PSM-DiD_e_l	(4) PSM-DiD_e_n	(5) PSM-DiD_l_k	(6) PSM-DiD_l_l	(7) PSM-DiD_l_n
DRcomputer	-0.008	-0.004	-0.005	0.000	0.004	-0.004
	(0.017)	(0.018)	(0.02)	(0.017)	(0.018)	(0.02)
DRfridge	-0.050	-0.048	-0.039	-0.086	-0.095	-0.102+
	(0.039)	(0.042)	(0.042)	(0.054)	(0.058)	(0.062)
DRiron	-0.024	-0.007	-0.010	-0.037	-0.026	-0.037
	(0.046)	(0.048)	(0.051)	(0.049)	(0.052)	(0.059)
DRmixer	-0.030	-0.028	-0.032	0.003	0.000	-0.014
	(0.025)	(0.028)	(0.03)	(0.028)	(0.03)	(0.033)
DRstereo	-0.013	-0.020	-0.024	-0.059	-0.048	-0.046
	(0.041)	(0.043)	(0.046)	(0.053)	(0.055)	(0.058)
DRstove	-0.013	-0.007	0.003	0.032	0.033	0.031
	(0.017)	(0.019)	(0.022)	(0.025)	(0.029)	(0.032)
DRvhsdvd	-0.088+	-0.095+	-0.098+	-0.176**	-0.208**	-0.222**
	(0.046)	(0.05)	(0.052)	(0.051)	(0.055)	(0.058)
DRwasher	-0.018	-0.015	-0.012	-0.045+	-0.047+	-0.059*
	(0.021)	(0.023)	(0.023)	(0.025)	(0.027)	(0.029)
SAN	0.162	0.196	0.164	0.067	0.095	0.146
	(0.124)	(0.142)	(0.162)	(0.166)	(0.177)	(0.198)
SANexcr_	0.052+	0.058	0.051	-0.013	-0.011	-0.003
	(0.028)	(0.036)	(0.041)	(0.038)	(0.045)	(0.052)
SANgarb_	0.053	0.058	0.044	-0.032	-0.029	-0.016
	(0.042)	(0.047)	(0.051)	(0.057)	(0.059)	(0.066)
SANvent_	0.073+	0.088+	0.076	-0.078	-0.071	-0.089
	(0.042)	(0.049)	(0.052)	(0.065)	(0.068)	(0.074)
SANwatr_	0.097*	0.11*	0.098+	0.021	0.044	0.053
	(0.044)	(0.047)	(0.054)	(0.049)	(0.053)	(0.06)
SANyard_	-0.029	-0.043	-0.039	-0.068	-0.075	-0.088
	(0.043)	(0.049)	(0.054)	(0.059)	(0.059)	(0.067)

**Significant to the 1%, *Significant to the 5%, + Significant to the 10%
In parenthesis: Bootstrapped St. Errors with 200 replications

The legend of each column is PSM-DiD_t_m. Where, t refers to the treatment group (e for early treated and l for lately treated); m refers to the matching method used (k for kernel matching, l for local linear regression and n for closest 5 neighbours).