The Impact of Three Mexican Nutritional Programs: The Case of Dif-Puebla

Daniel Zaga

Working Paper 09 | 2015

Centre for Finance and Development, Graduate Institute of International and Development Studies,
Maison de la Paix, Chemin Eugène-Rigot 2, 1211 Geneva 21, Switzerland

THE IMPACT OF THREE MEXICAN NUTRITIONAL PROGRAMS: THE CASE OF DIF-PUEBLA

Daniel Zaga Szenker*

September 2014

ABSTRACT

This paper presents an impact evaluation of three nutritional programs implemented in Puebla, Mexico, run by SEDIF, a social assistance institution. The present study uses both a propensity score matching and weighting in order to balance the treatment and the control groups in terms of observable characteristics, and to estimate, later on, the causal effect of the programs on different areas: food support, food orientation, education, and health. This investigation adds strong empirical evidence about the beneficial effects of nutritional programs on growth indicators (i.e. on anthropometric variables). In addition, it provides some evidence about the favorable impact of this kind of programs on food orientation outcomes, such as eating habit changes or diet diversity, variety, and quality. However, this study unveils only marginal effects on food security and detrimental effects on educational outcomes (specifically on student's marks). Finally, it does not provide conclusive effects on health.

Acknowledgements: I gratefully acknowledge the helpful comments of the UNDP Mexico Team (Dante Sánchez, Juan Francisco Pérez de la Torre, Juan Castaneda, Alma Lira, and Carlos Nandayapa). I am especially indebted to Erich Battistin and Jean-Louis Arcand for their invaluable comments and support. I also thank the Swiss Confederation and the Graduate Institute of Geneva for financial support.

Keywords: Nutritional programs, impact evaluation, anthropometrics, Mexico, Puebla.

JEL code: I12, O12, I20, D04, C31.

---

* PhD, Graduate Institute of Geneva.
Contact email: daniel.zaga@graduateinstitute.ch

I. Introduction

The purpose of this paper is to present the impact evaluation of the following three nutritional programs of the DIF Institution (*Desarrollo Integral de la Familia)* located in Puebla, Mexico: i) Hot School Breakfast (DEC by its acronym in Spanish), ii) Cold School Breakfast (DEF), and iii) Starting a Correct Nutrition (INC). DIF is a governmental agency in charge of conducting social assistance policies directed at strengthening family ties.

First, this study examines the methodological framework in which the impact evaluations are situated with the purpose of comparing the different evaluation tools currently available. The impact evaluation methods can be divided into two broad groups: the experimental group and the observational group. After analyzing the former, which it cannot be implemented due to the absence of an *ex-ante* random sample, the observational methods will be explored: i) matching and the propensity score; ii) instrumental variables; iii) RDD (Regression Discontinuity Design); iv) DID (difference-in-difference); and v) quantile regressions.

Second, a brief description of the three programs is offered, highlighting their main features and the variables to be evaluated. Based on the preceding sections, this study provides a justification for the quasi-experimental methods selected. Afterwards, some technical particularities of the estimations are pointed out: i) type of standard errors; ii) the need for fixed effects; and iii) some practical considerations for the implementation of the impact evaluation. Then, the results will be presented for each program.

Finally, the conclusions are presented, together with the policy implications and the recommendations for DIF-policy makers. At the same time, the results of the three programs are horizontally compared.

II. Methodological Framework

In the last few years, impact evaluations have received a remarkable attention in the public policy atmosphere. On the one hand, civil society participation in the public arena has led to a higher demand for more efficient public policies and for concrete and measurable results. On the other hand, governments have attempted to be perceived as more credible and accountable in order to increase public support.

These factors naturally derive in considering impact evaluations as a paramount tool for evaluating social programs, through which: i) policy makers are able to examine whether their programs generate the expected results; ii) government accountability is fostered; and iii) it could be unveiled which programs work and which is the magnitude of the impact attributable to the program (Khandker et al, 2010).

Impact evaluations represent a paradigm change with respect to the usual public policy analysis, which basically describes program budgets or only mentions the amount of beneficiaries covered by the program. By contrary, impact evaluations attempt to examine whether the program has reached its goal of enhancing wellbeing conditions, increasing income, improving education or decreasing diseases (Gertler et al, 2011).

Having said that, what do we mean by impact evaluations? What do they attempt to measure? Which are the difficulties for their implementation? Which methods can be utilized? These questions will be tackled in the following paragraphs of this section.

II.1. Some Precisions

Impact evaluation consists in determining the *causal effect* of an intervention on certain characteristics of a group of beneficiaries. Correspondingly, impact evaluation examines whether changes in some characteristics of the program's beneficiaries can be attributed to the intervention *per se* or to other factors. This paper will start by defining some key concepts of this literature: i) causal effect indicators; ii) definition of the counterfactual; iii) selection bias; iv) endogeneity bias; and v) selection of the counterfactual.

II.1.1. Causal Effect Indicators

In individual terms, the effect of an intervention is equivalent to the response variable for the treated unit (Y_{i1}) minus the same unit's variable value without intervention (Y_{i0}); i.e.,

$$\beta_i = Y_{i1} - Y_{i0} \tag{1}$$

In population terms, the **average treatment effect (ATE)** is given by the difference between the average of a treated group and the average of the same units had not received the intervention:

$$\beta = E[Y_1 - Y_0] = \sum_{i=1}^{N}(Y_{i1} - Y_{i0}) * 1/N = E[Y_1] - E[Y_0] \tag{2}$$

The ATE can be easily obtained from the basic econometric model of ordinary least squares, which departs from:

$$Y = \alpha + \beta T + \varepsilon \tag{3.1}$$

where T is a binary variable equal to 1 if under the program (and thus situated in the treatment group) or 0 if the intervention was not received (control group), α is a constant, β is the causal effect of the program and ε is the standard error with mean zero and constant variance. Then, if we calculate the expectation of Y given T=1, the expectation of Y given T=0, and their difference for achieving the ATE, the following is obtained:

$$E(Y|T=1) = \alpha + \beta T + E(\varepsilon|T=1) = \alpha + \beta;$$

$$E(Y|T=0) = \alpha + E(\varepsilon|T=0) = \alpha;$$

$$==> \beta = E(Y|T=1) - E(Y|T=0) = ATE \quad (3.2)$$

Generally, another indicator is used for measuring the causal effect: the **Average Treatment Effect on the Treated (ATT).** This indicator measures the average treatment effect given that the individual is participating in the program; i.e. the units from the treatment group are compared with similar units in the control group, instead of considering the whole population as in the ATE. In formal terms,

$$ATT = E[(Y_{i1} - Y_{i0})|T=1] = E[Y_{i1}|T=1]-E[Y_{i0}|T=1] \quad (4)$$

Having described the main causal effect indicators, now I turn to tackle the fundamental evaluation problem, which consists in the fact that each individual only faces one outcome -i.e. whether to participate in the program or not (Holland, 1986). In terms of equation 1, the beneficiary *i* observes Y_{i1} (outcome variable if participating), but cannot observe Y_{i0} (outcome variable without participation)[1]. This highlights the importance of the counterfactual term, which will be explained next.

II.1.2. Definition of the Counterfactual

A key question for the causal effect estimation can be summarized as: what would have happened had the individual not participated in the program? For obtaining the answer, Y_{i0} should be obtained for the beneficiary *i* (i.e. $Y_{i0}|T=1$). Since this cannot be observed, it turns into a missing data problem.

This term ($Y_0|T=1$), which is called the counterfactual, is estimated with the control group. Hence, finding an adequate control group becomes one of the main challenges in the impact evaluation arena. In practical terms, the treatment and the control group: i) should be, on average, statistically identical in the absence of the

[1] Seemingly, the individual *i* of the control group observes Y_{i0} (outcome variable if not participating), but not Y_{i1} (outcome variable if participating).

program; ii) should react in the same way if the program were implemented; and iii) could not be differentially exposed to other programs during the evaluation period (Gertler et al, 2011).

In experimental methods, this problem theoretically disappears since both groups have been randomly selected and thus their characteristics are statistically similar, obtaining unbiased estimations of the causal effect. The problem mostly arises when observational methods are used in the sense that various biases may be generated; particularly, the selection bias and endogeneity bias, which will be discussed in the following sub-sections.

II.1.3. Selection Bias

As already mentioned, a treatment and a control group should be selected, thereby a potential bias arises from the fact that the probability of being selected may be different for individuals of both groups. Hence, the challenge in impact evaluations is to select an adequate counterfactual; i.e. both groups should be identical in observable and non-observable terms, before the intervention.

What happens if the composition of the groups differs with respect to characteristics related to the outcome variable (Y)? For example, the beneficiaries may be more educated (observable characteristic) or motivated (unobservable characteristic) than the control group, thereby we may wrongly conclude that the program has beneficial results, whereas their real determinant is the differential composition of the groups. This situation is very common in impact evaluations, where the individuals: i) are self-selected into treatment; or ii) are selected on a geographical base. This pre-intervention difference in the composition of both groups, called *selection bias*, makes impossible to isolate the causal effect of the program.

If the difference between groups is based on *observable characteristics*, comparing similar units of both groups can solve the problem (either by matching or by the propensity score). However, the difficulty arises when both groups differ in *unobservable features*. There are various quasi-experimental methods for mitigating the selection bias problem that this paper will revise soon, such as RDD and DID.

In formal terms, if we were to evaluate the effect of a nutritional program on student's marks, we could examine the simple difference in the average marks between the treatment and control schools:

$$DIF = E[Y_1|T=\text{treated school}]-E[Y_0|T=\text{control school}] = E[Y_1|T=1]-E[Y_0|T=0]$$

If we add and subtract $E[Y_0|T=1]$, we obtain:

$$DIF = E[Y_1|T=1]-E[Y_0|T=0] + E[Y_0|T=1] - E[Y_0|T=1]$$

$$DIF = E[Y_1|T=1] - E[Y_0|T=1] + E[Y_0|T=1] -E[Y_0|T=0]$$

$$DIF = E\ [(Y_{i1} -Y_{i0})|T=1] + \{E[Y_0|T=1] -E[Y_0|T=0]\}$$

$$DIF = ATT + \{E[Y_0|T=1] -E[Y_0|T=0]\} \tag{5}$$

Equation 5 shows that the difference between groups is equivalent to the ATT plus a term, which is equal to the selection bias. This last term refers to the difference between groups had the program not been implemented. The purpose of every impact evaluation, thus, is to identify situations where we can assume that the selection bias is inexistent or where we can find strategies to correct for it (Duflo et al, 2006).

II.1.4. Endogeneity

Impact evaluations aim at comparing the outcome variable *Y* between individuals of the treatment and the control groups. By adding the control variables *X,* equation 3.1 becomes:

$$Y_i = \alpha X_i + \beta T_i + \varepsilon_i \tag{6}$$

One of the basic assumptions of the OLS method, which generates efficient and unbiased estimates, is that the explanatory variables (*X* and *T*) cannot be associated with *e*; i.e. they should be exogenous.

However, in the impact evaluation context, there could be unobservable variables in ε correlated with the probability of participation (*T)*, which, in turn, determines the outcome variable *Y*. This **problem of omitted variables** is called *endogeneity* (Wooldridge, 2002), where OLS estimates turn biased and inconsistent. In formal terms:

$$\mathrm{cov}(T,e) \neq 0 \tag{7}$$

Endogeneity bias may arise if: i) the selection rules are not clear; ii) program participation is not compulsory; or iii) individuals find the way to skip their assigned status. Thus, *T* becomes endogenous. There are two solutions for this problem.

First, in a panel data context, this problem is solved by adding **individual fixed effects**, assuming that unobservable characteristics are time-invariant. Briefly speaking, this model transforms each variable in its difference with respect to the average over time for each individual, and OLS is applied later over the transformed variables. In this way, the time-invariant variables (both observed and unobserved) are wiped-out, thus the causal effect is cleansed from individual heterogeneity[2].

Second, when data is structured in a cross-section manner and the unobservable characteristics vary over time, the **instrumental variable approach** may be implemented. This strategy will be analyzed in the following sections.

---

[2] When there are two periods of time, the results from this model are equivalent to DID results, as it will be analyzed soon.

II.1.5. Selection of the Counterfactual

The program causal effect is obtained by comparing the outcome variable between the treatment and the control group (counterfactual), both with statistically similar characteristics. Therefore, the selection of the counterfactual plays a crucial role in the impact evaluation scenario.

Various methods may be used to create valid control groups. Though this will be explained in detail in the next sub-sections, it is interesting to mention two simple or *naïve* methods that are clearly biased, with the purpose of unveiling common estimation errors. Gertler et al (2011) define these methods as: i) *before-and-after* comparison; and ii) *with-and-without* comparison.

The **before-and-after method** compares the outcome variable for the beneficiary after (Y|T=1) and before (Y|T=0) the treatment, assuming that the outcome variable would be constant had the beneficiary not exposed to the treatment. In a more sophisticated scheme, control variables may be included in the estimations for controlling for observables. However, considering a control group is not included into the evaluation, beneficiaries' unobservables may be driving the results, thus leading to debatable conclusions.

The **with-and-without comparison** method differentiates a treatment and control group in a pretty unsophisticated way. For example, the government may offer a nutritional program to the whole spectrum of schools in a specific community; i.e. this method would compare the schools that voluntary accepted to be part of the program with all the rest of schools. The problem of this strategy can be easily observed in equation 5, where selection bias may be rooted in both observable and unobservable variables.

Having in mind the clear drawbacks of these two simplistic methodologies, we move on to the *main methods for creating the counterfactual* in impact evaluations. These methods depart from different *set of rules* for the selection process of both groups. Briefly speaking, the selection process (i.e. the determinants of T for each i) depends on three factors: i) observable variables (X); ii) unobservable variables (U);

or iii) a random sample (*Z*); i.e. the selection process may be summarized as: T=T(X,U,Z).

First, we will analyze the randomized controlled trials (RCTs), where T=T(Z). Afterwards, we will continue with the quasi-experimental methods in the following order: i) matching and the propensity score where T=T(X,U) and *U* is independent from *Y*, given *X*; ii) instrumental variables where T=T(X,U,Z); iii) regression discontinuity design (RDD) where T=T(X); iv) difference-in-difference (DID) where T=T(X,U) and *U* is independent from the variation of *Y* over time; and v) quantile regressions where *T* may be a function of any of those factors.

It is important to point out that these methods can be combined. For example, it is of usual practice to use RCTs or the propensity score matching together with the DID method in order to generate more robust results.

II.2. Different Methodologies

The experimental methods (or RCTs) are considered the "golden rule" in the evaluation literature, since, if well designed and implemented, it may lead to unbiased results (Sefton et al, 2002). Nevertheless, different circumstances may derive in the need for observational methods, where several biases can arise. This fact has given birth to a large debate about the suitability of each method.

Lalonde (1986) has initiated this debate. He estimated the effect of an employment program in which individuals were randomly selected into the treatment and control group. He compared these results with those obtained from non-experimental methods and concluded that the different methodologies provide divergent results. After that influential paper, various authors carried out similar comparisons and some of them challenge Lalonde's results. For example, Glazerman, Levy and Myers (2003) compare both methodologies by the analysis of twelve programs, finding similar results across both methodologies in *only* some occasions.

It is worth mentioning, however, that several differences exist within the observational methods. As we will revise soon, RDD is considered as the most precise strategy in quasi-experimental methods since its results are unbiased under certain circumstances (Cook, Shadish and Wong, 2006; DiNardo and Lee, 2010; Buddelmeyer and Skoufias, 2004).

Summing up, there is no single ideal method in impact evaluations. The selection of the most appropriate tool would depend on the economic model, data availability, and the questions to be solved (Heckman, Lalonde and Smith, 1999).

II.2.1. Experimental Methods

In the impact evaluation context, if treatment and control group characteristics were not associated with the outcome variable, the optimal "laboratory" solution would be to randomly assign the eligible units to each group. Under this context, each unit has the same probability to be selected to treatment, considering a large number of potential units to apply randomization.

In formal terms, the set of rules of this method is represented by T=T(Z), thus selection bias is mitigated. This is equivalent to say that both groups are balanced by observables (*X*) and unobservables (*U*), as a consequence of the selection process (DiNardo and Lee, 2010). In other words, $E[Y_0|T=1]$ is equal to $E[Y_0|T=0]$ in equation 5, since the likelihood of participation is equal for every unit and, thereby, the difference in the outcome variable between the groups (i.e. DIF) is the ATT, equivalent to ATE.

The experimental methods also offer some advantages for program administrators, because they cannot be accused of favoring some individuals, taking into account that the selection process is random and, therefore, difficult to manipulate (Gertler et al, 2011).

On the contrary, this method presents some disadvantages. First, there is an ethical issue, since not all the individuals participating in the evaluation receive a benefit from the program -see Dobash et al (1999) for an example. Second, this method can be considerably expensive with respect to quasi-experiments -e.g. see Olken (2005). Third, it is not always viable to perform an RCT. In particular, Jalan and Ravallion (2003) highlight that some programs need to be quickly implemented as a response to an economic crisis. In those cases, a pre-intervention randomization is not feasible. Lastly, and related with the last point, impact evaluations carried out with experimental methods may take a long time, thus they are not necessarily policy-oriented.

Finally, it is important to differentiate two concepts: **internal and external validity**. The former refers to the potential bias in the causal effect estimation[3], whereas the latter is related to the fact that the impact found may be generalized to the whole eligible population. Since RCTs are unbiased as a consequence of the randomization process, this method is internally valid. This is an important characteristic to consider when comparing it with the other methods. The external validity of RCTs depends on the eligible population facing randomization, which is not necessarily the whole population; e.g. randomization may be conditioned by certain observable variables, such as vulnerability levels or individual income.

For assuring external validity, randomization can be performed in two steps: i) randomization is done over the whole eligible population in a representative way (assuring external validity); and ii) over that sample, randomization is applied again for determining the units assigned to each group (keeping internal validity).

II.2.2. Quasi-Experimental or Observational Methods

When the analyst cannot manipulate the selection process, or whenever it is unethical to do so, other strategies may be found to carry out an impact evaluation. The purpose of the quasi-experimental methods is to find the most similar counterfactual to the one

[3] DiNardo and Lee (2010) define internal validity as the degree of correspondence between what is known about the selection process and the statistical model of the analyst.

obtained from RCTs. In this part of the research, the following methods will be presented: i) matching and the propensity score; ii) instrumental variables; iii) RDD; iv) DID; and v) quantile regressions.

II.2.2.1. Matching and the Propensity Score

The matching procedure allows the analyst to design a counterfactual based on observable characteristics. Individuals of both groups should be similar in terms of observational variables not affected by the program (baseline data or time-invariant conditions). In practical terms, each beneficiary should be matched with a non-beneficiary and, afterwards, the difference of the average of both groups is taken to obtain the causal effect.

The key condition of this method is that unobservable features associated with the outcome variable should be statistically similar between groups. Otherwise, these estimations would be biased[4]. Therefore, under this method, the analyst should acquire a large number of observable characteristics (X), with the purpose of reducing the potential selection bias. Practically speaking, it is difficult for the analyst to match a great quantity of individuals with the same X's. This problem has been called the "curse of dimensionality".

Thus, the propensity score (PS) appears as a natural replacement for matching. The PS is a *balancing score*, since the distribution of the observable characteristics is similar between groups, given the PS (Rosenbaum and Rubin, 1983), reducing the multidimensional problem of matching to a one-dimensional. In formal terms, the PS creates a counterfactual based on the likelihood of participation, given the observable variables, where the selection process is T=T(X,U) and *U* is independent from *Y,* given *X*.

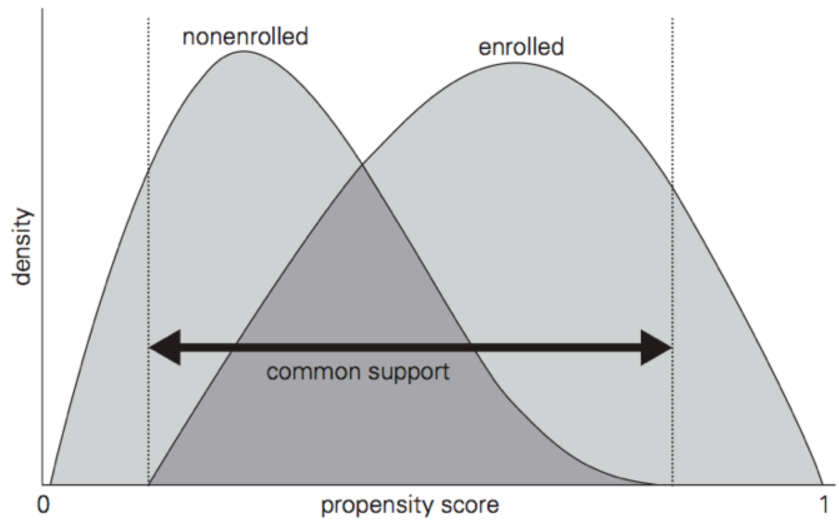
Rosenbaum and Rubin (1983), and Heckman, Lalonde and Smith (1999) sustain that some assumptions should be done under the PS calculation. First, the treatment on individual *i* should not affect the individual *j* (i.e. SUTVA: *Stable Unit-*

[4] That is way matching is usually combined with other methods of evaluation, such as DID.

Treatment Value Assumption). Second, the outcome variable *Y* should be independent from *T*, given *X*, what is called the *Ignorable Treatment Assignment* or the *Conditional Independence Assumption*. In technical terminologies, $(Y_{i1}, Y_{i0}) \perp T_i \mid P(X_i)$, where $P(X_i)=P(D=1|X_i)$[5]. Lastly, the common support assumption should be conformed; i.e. treated units should have comparable units in the distribution of the PS; i.e. $0 < P(T_i = 1|X_i) < 1$. Rosenbaum and Rubin (1983) call *strong ignorability* when these last two assumptions are achieved.

The last assumption generally leads to the "problem of common support", which arises when a large proportion of individuals need to be eliminated from the analysis for assuring group comparability. This is a usual problem, since the treated group may not contain individuals with low PS; seemingly, the control group may not include units with high PS. These two possibilities can be visualized in the tails of the PS distributions in Figure 1. In sum, external validity is rather problematic under this methodology of evaluation.

FIGURE 1: The Common Support Region



Source: Gertler et al (2011).

In addition, the internal validity of this method would be only satisfied if the three preceding assumptions were valid (Khandker et al, 2010).

[5] The practical problem of these first two assumptions is that they cannot be tested.

Briefly speaking, there are three methods under the umbrella of the PS. The first one is the use of the PS as a covariate. The problem of this methodology is that it is based on the strong assumption that the relationship between the PS and the outcome variable has been correctly modeled (Austin, 2011). Thus, in the following sub-sections, we are going to analyze in detail the other two options: i) the Propensity Score Matching or PSM*;* and ii) the Propensity Score Weighting or PSW.

II.2.2.1.a. PSM

After obtaining the PS of each unit, the PSM calculates the causal effect as the difference in the outcome variable between the treatment and the control group, weighted by the distribution of the PS, based on a matching technique. The PSM may determine both the ATE and the ATT. When the method is not externally valid, because some units are not included in the analysis, only the ATT can be calculated. In a cross-section data structure, Smith and Todd (1995) define the ATT as:

$$\text{ATT} = \frac{1}{Nt}\left[\sum_{i \in T} Y_{\text{i}}^{\text{T}} - \text{w}(_{\text{i,j}}) Y_{\text{j}}^{\text{C}}\right] \tag{8}$$

where N_t is the total number of beneficiaries *i*, while $w(_{i,j})$ is the weighting used over the control group.

In brief, the PSM estimations are valid if the control and treatment groups: i) have the same distribution of unobservable characteristics (which cannot be tested); ii) have the same distribution of observable characteristics; iii) answer the same questionnaire; and iv) reside in the same economic environment, thereby facing the same economic incentives that may define their participation into the program (Jalan and Ravallion, 2003; Heckman, Ichimura, and Todd, 1997; Ravallion, 2008).

There is a debate over the number of variables to include in the model in order to determine the PS. Heckman, Ichimura, and Todd (1997) suggest that the omission of relevant variables may significantly increase the bias of the results, implying that all pre-treatment and time-invariant variables should be added into the model, if they are related to T and Y. On the other hand, Bryson, Dorsett and Purdon (2002) conclude that models should not be over-parameterized, since the inclusion of non-relevant variables may exacerbate the common support problem, and may increase the variance of the results. Finally, Rubin and Thomas (1996) propose a practical advice: if there were uncertainty whether to include or not a variable, it would be better to leave it in the model.

The last topic related to the PSM is associated with the fact that the likelihood of finding exactly the same PS between both groups is zero, given that Pr(T=1|X) is a continuous variable. Therefore, a matching criterion should be selected. We will revise four techniques: *Stratification Matching*, *Nearest-Neighbor Matching*, *Radius Matching*, and *Kernel Matching*. There is no *ex ante* preferred matching method; rather, this would be selected depending on the particular situation of each evaluation.

The *Stratification Matching* divides the distribution of the PS in blocks, having in mind that each of them should have individuals from both groups with a similar PS average. Afterwards, ATT is calculated for each block, and then, the final ATT is obtained as the ATT mean among blocks, weighted by the share of participants in each interval. The problem of this technique is that it eliminates those individuals without match, thereby putting into question its internal and external validity.

The *Nearest-Neighbor Matching* (or *NN Matching*) is a technique that matches individuals from the treatment group with those individuals from the other group with the closest PS. The non-participants may become a unique match (*without replacement*) or may be matched with more than one participant (*with replacement*). The advantage of this technique is that all the units can be matched, depending on the selected range. However, the drawback is that it may match individuals with a considerable long distance between their PS, hence leading to inaccurate estimations of the causal effects.

The *Radius or Caliper Matching* consists in matching those individuals located in the same PS radius. The larger the radius, the less precise the results. The shorter the radius, more units should be eliminated from the analysis.

Finally, *Kernel Matching* is a non-parametric method that matches each individual of the treatment group with the weighted average of all the units of the control group. Weights are inversely proportional to the distance between the PS of the matched units. The higher the distance between the PS, the less weight for the ATT calculation.

As a summary of this sub-section, the researcher should follow the next steps in order to carry out the PSM: i) to estimate the PS (which can be simplified to a *logit* or *probit* estimation if no unit were discarded for keeping the common support); ii) to define the common support region and to perform the balancing tests; iii) to decide the matching technique; and iv) to estimate the causal effect (Khandker et al, 2010)[6].

II.2.2.1.b. PSW

We have seen in equations 3.1. and 3.2. how to obtain β = ATE from Y = $\alpha + \beta T + \varepsilon$. Under the same logic, the PS can be utilized as a weight for the calculation of the causal effect, with the purpose of balancing the treatment and the control groups. Replacing Y_0 for α, we depart from:

$$Y = Y_0 + \beta T + \varepsilon \tag{9}$$

Now, we replace β for Y_1-Y_0, based on equation 1; weighting the treated group by 1/PS and the control group by 1/(1-PS), the following equation is obtained:

$$Y = \frac{Y_1 T}{PS} + \frac{Y_0(1-T)}{(1-PS)} \tag{9.1}$$

We take expectations for both groups and their differences:

$$E(Y|T=1) = \frac{Y_1 T}{PS};$$

[6] Caliendo and Kopeinig (2005) offer a more comprehensive version of the PSM steps.

$$E(Y|T=0) = \frac{Y_0(1-T)}{(1-PS)};$$

==> $$E(Y|T=1) - E(Y|T=0) = E[\frac{Y_1 T}{PS} - \frac{Y_0(1-T)}{(1-PS)}] = ATE = \beta$$

This estimator weights both groups to a common distribution of observable variables; i.e. the marginal distribution of X for the whole population. Hence, the ATE can be estimated as a weighted average, through the inverse probability of treatment weights (IPTW). These weights balance, in expected terms, the distribution of observable variables between groups. In addition, PSW determines consistent, unbiased (Imbens, 2004), and, under some circumstances, efficient estimators.

Hirano and Imbens (2001) suggest two variations for the ATT estimation with PSW. First, weights could change to 1 for the treated group and to PS/(1-PS) for the control group[7]. Second, they propose to add the interaction between the treatment variable (*T*) and the difference between each control variable (X) and its mean for the treated units ($\bar{X}$) -i.e. (X-$\bar{X}$)*T_1- in order to control for non-additive associations.

One possible inconvenient of the PSW is that weights can be unstable in the tails of the distribution of T, increasing the variation of the estimated causal effect (Austin, 2011). As a potential solution, those individuals can be eliminated in order to find an appropriate common support, avoiding the inclusion of outliers.

A particular PSW estimator is the parametric double-robust (DR) estimator, applied in both the PS (first step) and the causal effect (second step) estimations, conceived by Robins, Rotnitzky and Zhao (1995)[8]. The name of this estimator is related to the large sample property, which determines that $\hat{\beta}$ is a consistent estimator of β if either the first or the second step is correctly specified (Imbens, 2004). Nevertheless, the DR: i) is no longer efficient if only one step is correctly specified

[7] These weights are not the only ones used in PSW (Imbens, 2004); however, they are the most frequent.

[8] Robins et al (1995) show that the use of the estimated weight (*vis-á-vis* the real weight) increases the efficiency of the estimators in the parametric models.

(Lunceford and Davidian, 2004); ii) should include the same variables in the two estimation steps; iii) can only be implemented in large samples; and iv) does not provide any hint about how to know if the models are correctly specified.

II.2.2.1.c. Comparison of the PS Methods

Austin (2011) suggests that the PSM (except with *Stratification Matching*) and the PSW offer a better sample selection correction than using the PS as a covariate. Rubin (2004) highlights that PSW and the PS as a covariate are more sensitive than the PSM. However, Rubin (2001) points out that the PSM and the PSW are more desirable than the PS as a covariate, since they differentiate between the design and the analysis of the study. That is, first, the PSM and the PSW estimate the PS in order to satisfy the balancing conditions; afterwards, they estimate the causal effect; however, in the remaining case, the same regression includes *Y, T* and the *PS* estimated, thus the analyst may be tempted to find the PS that leads to the *Y* that he or she expects.

In sum, the debate about the most suitable PS estimator is still on going without conclusive agreements (Imbens, 2004). However, there is some consensus about the potential drawbacks of using the PS as a covariate. In any case, the selection of the appropriate PS method would depend on the specific circumstances of each research.

II.2.2.2. Instrumental Variables

As commented in sub-section II.1.4., *T* is generally endogenous, specifically in quasi-experimental methods, thus OLS estimators become biased and inconsistent. The instrumental variable (henceforth IV) method, which may solve the problem, seeks a variable *Z* (the IV) that fulfills the following requirements: i) it should be significantly associated with the participation variable *T*; and ii) the only association between *Z* and the outcome variable *Y* should be channeled through *T*; i.e., *Z* cannot be associated with *Y* through the error term (which comprises all the unobservables that the model is not able to capture) of the structural equation, what is called the *exclusion restriction*.

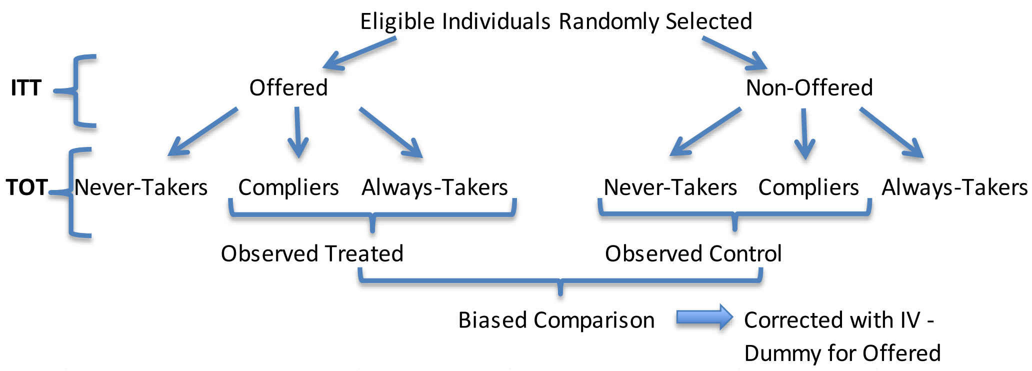
In the impact evaluation literature, the random selection of individuals -in the first step- has been generally used as the most preferred IV. Since not all the individuals comply with their assigned state (whether to participate or not), three types of individuals may arise: i) the *compliers* (those accepting their assigned state); ii) the *never-takers* (those never participating); and iii) the *always-takers* (those always located in the treatment group).

Hence, first, the causal effect between the selected and the non-selected individuals is estimated -called the **ITT** or the *intent to treat estimate*. This estimator is important for policy-making, since individuals may be offered, but not forced, to accept their assigned state (Gertler et al, 2011). However, if the purpose of the evaluation would be to examine the effect on those individuals who effectively receive the program, another indicator would be pursued: the **TOT** or *treatment on the treated*. This estimator is obtained comparing the groups that effectively participate and non-participate (note that Z and T differ). Since a direct comparison between the observed groups would determine a biased estimation (since T is endogenous), Z (the eligibility criterion, which is random) is used as the IV for T (the effective participation), which in turn, determines Y.

It is important to highlight that Z fulfills the *exclusion restriction*, since the selection process is random and, thus, it could not be systematically associated with unobservables determining the outcome variable. Figure 2 illustrates this evaluation methodology, called *Random Offering*.

In sum, the set of rules that determines the selection process is represented by T=T(X,Z,U), where T differs from Z, and U is correlated with Y_0, so sample selection arises.

FIGURE 2: **Random Offering**



Source: Own elaboration.

In formal terms, the causal effect of the program β_{IV} can be calculated as the ratio between cov(Y_i, Z_i) and cov(T_i, Z_i), which can be intuitively seen as the relationship between *Y* and *Z*, minus the portion of Z that explains T. Starting from $Y_i = \beta T_i + e_i$, and considering that the *exclusion restriction* is complied (i.e. cov(Z,e)=0), we obtain:

$$\mathrm{cov}(Y_i, Z_i) = \mathrm{cov}[(\beta T_i + e_i), Z_i] = \beta \mathrm{cov}(T_i, Z_i)$$

$$\Rightarrow \frac{\mathrm{cov(Yi,Zi)}}{\mathrm{cov(Ti,Zi)}} = \beta \qquad (10)$$

The coefficient β determines the causal effect of the program for the compliers. Angrist and Imbens (1994) describes this result as a local ATE (LATE), defined as the average treatment effect for those individuals with a treatment status affected by a change in exogenous regressors that satisfy the exclusion restriction. From the estimation of β, the ITT estimator can be easily obtained; i.e. E(Y|Z=1) - E(Y|Z=0).

Unlike other quasi-experimental methods (such as the PSM or the DID), an important benefit of this strategy is that it does not need to make assumptions over sample selection. This means that by finding a strong instrument (i.e. *Z* highly associated with *T*) not related to unobservable characteristics determining Y, the endogeneity problem is mitigated. Therefore, if the IV requirements are fulfilled, the causal effect is internally valid; however, the external validity is reduced to only the eligible population (ITT) or the compliers (TOT).

Other instrumental variables have been used in the impact evaluation literature. Arcand and Bassole (2006) used an IV estimation, among other methodologies, in their impact evaluation of PNUR -a community driven development program- in Senegal. They used community leader opinions and projections (as a proxy for their commitment with the community) as IVs, with the presumption that those communities with more active and participative government heads would have a higher probability of participating in the program. In another study, Glewwe and Jacoby (1995) examine the effect of nutrition and health on education in Ghana. Their identification strategy consists in using the distance from health facilities and mother weights as IVs for child health. This study reveals the difficulty in finding a valid instrument that satisfies the *exclusion restriction*, considering that it is highly unlikely that these IVs were unrelated to unobservables associated with education.

Summing up, the IV method is a valid tool for determining a program causal effect when the IV requirements are complied. Nevertheless, its results are not externally valid and its implementation is highly dependent on data availability.

II.2.2.3. Regression Discontinuity Design

The regression discontinuity design (RDD) is considered as the most robust strategy within the observational methods, since its causal inference is the most closely linked to randomization (Cook, Shadish and Wong, 2006; DiNardo and Lee, 2010; Buddelmeyer and Skoufias, 2004).

RDD is an estimation strategy where treatment is realized when an observed, forcing or running variable S exceeds a known threshold (s*); i.e., the selection process is given by T=T(X), where T=1 if $S \geq s^*$ and $S \in X$. One condition of this strategy is that the probability of treatment assignment should be a discontinuous function of one or more variables.

The theoretical background of this strategy sustains that individuals surrounding the threshold s* have very similar characteristics. Accordingly, a treatment and a control group can be identified if individuals were located "just" over/under the threshold[9]. In this way, selection bias is mitigated. This characteristic of RDDs is similar to a local randomization; therefore the estimate, which is called LATE or local ATE, is internally valid. However, RDDs are only externally valid for sub-populations close to the threshold; i.e. when S tends to s* (Imbens and Lemieux, 2008).

The fact that this method is internally valid represents a substantial advantage for RDDs compared to the rest of the observational methods, which generally require that the unobservable characteristics be independent from program participation. However, it is completely different if independency is an assumption of the method, rather than a consequence of the process of data generation (Lee, 2008).

As regards the causal effect indicator, Imbens and Lemieux (2009) define the ATE for a sharp RDD as:

$$\text{ATE} = \lim_{S\downarrow s*} \text{E}\,[\text{Yi}|\text{Si} = \text{s}] - \lim_{S\uparrow s*} \text{E}\,[\text{Yi}|\text{Si} = \text{s}] = \text{E}\,[\text{Yi}(1) - \text{Yi}(0)|\text{Si} = s\,*] \qquad (11)$$

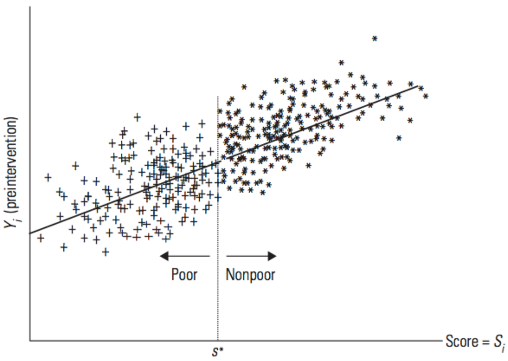
Equation 11 shows that ATE is calculated as the difference in the average of the outcome variable between those individuals just over the threshold and those just under it. This definition reveals some uncertainty about the distance from S to s* since the shorter the distance, the higher similarity among individuals, but the smaller the sample size and the power of the estimation -which is zero in the limit, when S=s* (Lee and Lemieux, 2010).

The effect of an intervention can be easily observed through the following figures, considering an example where the vulnerable (and treated) group is defined if

[9] This is not an assumption; by contrary, this can be tested.

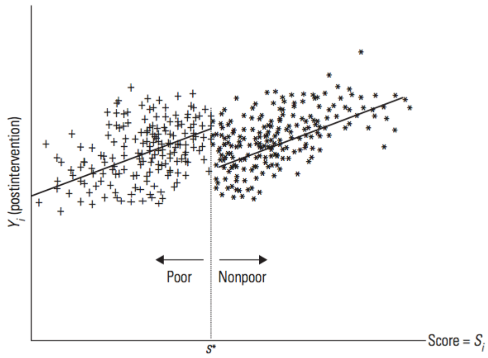
S<s*. An intervention may consist in food orientation talks for those households located under the poverty threshold where the outcome variable is household diet diversity (where *S* is the poverty level, *s** is the poverty threshold and *Y* is diet diversity). Figure 3 illustrates the pre-intervention context where the relation between *S* and *Y* linearly and constantly grows, whereas Figure 4 clearly reflects: i) the discontinuity in their post-treatment relationship; and ii) the jump of the outcome variable in the treatment group, equivalent to the causal effect.

FIGURE 3: Pre-Intervention



Source: Khandker et al, 2010.

FIGURE 4: Post-Intervention



Source: Khandker et al, 2010.

It is important to notice that the discontinuity previously shown was manifested under a linear relationship between *S* and *Y*. However, the association between these variables may follow a more complex functional form (e.g. cuadratic). Hence, the analyst should examine the most suitable functional form that reflects the real nature of the data (Gertler et al, 2011).

Two distinctive strategies may be differentiated within the regression discontinuity design. First, the **sharp RDD** is contextualized when the treatment status deterministically follows the selection rule T=1 if S ≥s*. Yet, if individuals would manage to change from their assigned group (i.e. *non-compliance*), a **fuzzy RDD** should be implemented, where *Z* differs from *T*. This happens very frequently in social programs where *S* determines eligibility, but not everyone accepts the rule. In this last strategy, *Z* is an IV for *T* and a Wald estimator is obtained. The difference between a sharp and a fuzzy RDD is equal to the difference between randomized assignment and offering in the context of experimental methods (Lee and Lemiux, 2010).

In brief, taking into account that RDD is considered the "cousin" of randomization (Lee and Lemieux, 2010), if there were a forcing variable determining a discontinuity in the selection process, RDD would represent the most attractive impact evaluation strategy.

II.2.2.4. Difference-in-Difference

The difference-in-difference (DID) method compares changes over time (between pre- and post-intervention) in the outcome variable between the treatment and the control group. The implementation of DID requires panel data (at least two observations per individual[10]) or repetitive cross-section data if the composition of each group is relatively stable over time.

It is not a pre-requisite to specify the set of rules for the selection process of units. That means that both groups are selected without any explicit set of rules. That is why DID is frequently combined with randomization or the propensity score. It should be noticed that the causal effect within the last methods has been analyzed, so far, with a simple difference that only requires cross-section data. Through the combination of different methods, the robustness of the results increases to a large extent.

DID represents a combination of the two previously analyzed simplistic methods, the **before-and-after comparison** and the **with-and-without comparison**. Having two points in time (t=1 and t=0) and two groups (T=1 and T=0), Y_t^T can be obtained and, consequently, ATT can be calculated through:

$$DID = ATT = E[(Y_1^T - Y_0^T)|T=1] - E[(Y_1^C - Y_0^C)|T=0] \quad (12)$$ [11]

[10] As a result of following the same individual over time, attrition bias may arise.
[11] It can be easily shown that DID equals ATT, and that under certain circumstances (such as randomization), equals ATE.

DID estimates may also be obtained in the typical econometric context, where temporal shocks affecting both groups (t) and unobservable characteristics of each group (T) are controlled for, through the following regression:

$$Y_{it}=c + \alpha t + \beta T_{i1} + \gamma(t * T_{i1}) + \varepsilon_{it} \quad (13.a)$$

DID is obtained as the difference in expectations of each group between post- and pre-treatment status, equal to γ (the interaction term between t and T):

$$E[(Y_1^T - Y_0^T)|T=1] = (c + \alpha + \beta + \gamma) - (c + \beta) = \alpha + \gamma \quad (13.b)$$

$$E[(Y_1^C - Y_0^C)|T=0] = (c + \alpha) - c = \alpha \quad (13.c)$$

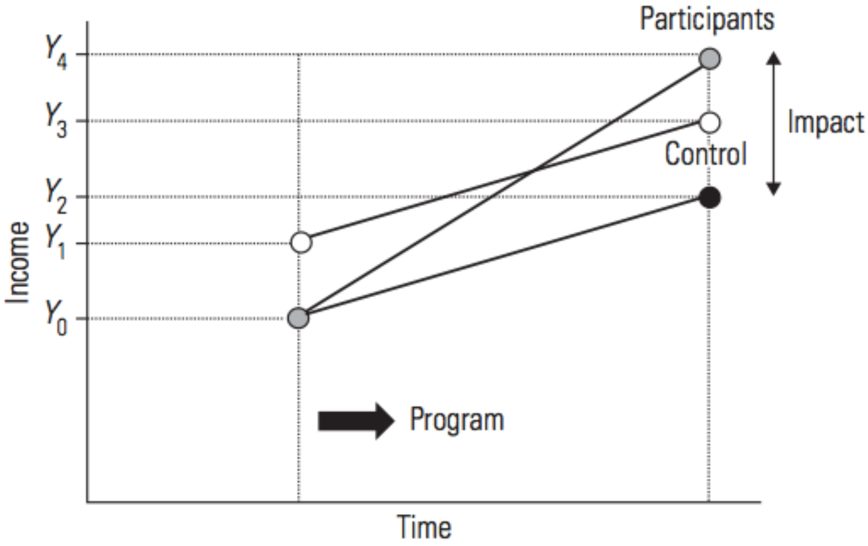
$$DID = E[(Y_1^T - Y_0^T)|T=1] - E[(Y_1^C - Y_0^C)|T=0] = \gamma \quad (13.d)$$

What is the main benefit of this evaluation strategy? By differencing the variables over time, individual time-invariant characteristics are wiped-out, not only observables (such as parents' education) but also unobservables (such as motivation or ability). However, time-varying characteristics cannot be balanced between groups. Thus, an important assumption of DID is that these differences do not exit, thereby both groups would be equal in the absence of the program[12]. This is formally called as the *Parallel-Trend Assumption*, which can be easily tested through another DID estimation with two pre-treatment periods.

Even though it does not present a precise set of rules, the DID selection process is formally considered as T=T(X,U) -i.e. not depending on *Z*- and *U* is assumed to be correlated with *Y* but uncorrelated with ΔY, complying with the Parallel-Trend Assumption. This is a key assumption for estimating DID, which is equal to $(Y_4 - Y_0) - (Y_3 - Y_1)$ in Figure 5. This assumes that $(Y_3 - Y_2) = (Y_1 - Y_0)$, and thus DID = $(Y_4 - Y_2)$.

[12] Ravallion (2008) sustains that this assumption is hardly fulfilled in the poverty program context of developing countries.

<u>FIGURE 5</u>: DID Calculation



<u>Source</u>: Khandker et al (2010).

Finally, it is relevant to mention that a two-period DID can be generalized to a fixed-effects model with panel data and numerous periods. This model departs from the premise that T_{it} is correlated with the unobservable individual time-invariant heterogeneity (ν_i); i.e., *T* is endogenous, as previously defined. Hence, equation 13.a. is revised to:

$$Y_{it} = \gamma T_{it} + \beta X_{it} + \nu_i + e_{it} \tag{14}$$

Differencing each variable over time, we obtain:

$$\Delta Y_{it} = \gamma \Delta T_{it} + \beta \Delta X_{it} + \Delta e_{it} \tag{15}$$[13]

After ν_i is removed, OLS can be applied to equation 15, obtaining the DID estimate. In a context of more than two periods of time, DID differs from the results obtained from a fixed effects model (Khandker et al, 2010).

[13] This equation represents the *first-differencing model*, equivalent to the *fixed effects model* in a panel of two periods. The *fixed effects model* takes the difference of each variable with respect to the average over time for each individual, thus individual time-invariant heterogeneity is eliminated.

II.2.2.5. Quantile Regression

So far, the methods already analyzed provide estimates of the average effect of the intervention. However, it can be very useful to figure out the effect of a program on different points of the outcome variable distribution, since the causal effect is not necessarily the same along different individuals (Buchinsky, 1998).

For example, the purpose of a program that supplies books in schools may be to increase not only students' marks averages, but also those of a particular quantile in the distribution. Even more, the program may attempt to compare the effect on different quantiles, more relevant for policy implications.

We have attempted, so far, to obtain the causal effect β that minimizes the mean square error of the estimation through OLS. That is, from $Y_i = \alpha + \beta X_i + \varepsilon_i$, where $T \in X$, we obtained $E(Yi|Xi) = \beta Xi$ and, thereby, $\partial E(Y_i|X_i)/\partial X_i = \beta$, equivalent to the causal effect.

In this case, however, we will get $Q\tau(Y_i|X_i) = \beta\tau X_i$, the conditional quantile of Y, given X, and thus $\partial Q\tau(Y_i|X_i)/\partial X_i = \beta\tau$, equivalent to the causal effect at different values of the distribution of Y. In more technical terms, this is equivalent to the partial derivative of the conditional quantile of Y with respect to X. In sum, the quantile regression is obtained by minimizing the absolute deviations with asymmetric weights (Koenker and Bassett, 1978):

$$min\, \beta \left[\sum_{t\epsilon\{t:y_t \geq x_t b\}} \tau|y_t - x_t b| + \sum_{t\epsilon\{t:y_t < x_t b\}} (1-\tau)|y_t - x_t b| \right] \tag{16}$$

In the impact evaluation context using quantile regressions, the relevant causal effect indicator becomes the **QTE** (quantile treatment effect). This is equivalent to the difference in the outcome variable Y between the treatment and the control group, located in the quantile τ from Y, if the units have been *randomly* selected:

$$QTE\,(\tau) = Y^t\,(\tau) - Y^c\,(\tau) \tag{17}$$

Yet, this equation should not be applied in quasi-experimental methods, since, among other reasons, it cannot be assured that the counterfactual of the treated individual *i* is located in the same quantile of the control group. In more technical terms, this occurs because the identification of QTE lies on the marginal distribution of Y_1 and Y_0, which is not achieved in observational methods. Even more, unlike the ATE where the expected value is a linear operator, and thus, $E[Y_{i1}-Y_{i0}|X_i] = E[Y_{i1}|X_i] - E[Y_{i0}|X_i]$, the functional of the difference in the conditional quantiles is not equal to the difference in the functionals of each group and each quantile (Heckman, Smith and Clements, 1997); i.e.:

$$Q\tau(Y_{i1} - Y_{i0}|X_i) \neq Q\tau(Y_{i1}|X_i) - Q\tau(Y_{i0}|X_i) \qquad (18)$$

In sum, *T* is endogenous in observational methods, thus conventional quantile regressions are inconsistent and, therefore, inappropriate for estimating the causal effect (Fröhlich and Melly, 2008). Consequently, different strategies have been proposed to overcome this issue.

First, some authors have tried with instrumental variables. Abadie, Angrist and Imbens (2002) implemented an IV estimation under conditional QTE with respect to X in order to solve the endogeneity bias of *T*, obtaining QTE for the compliers. Fröhlich and Melly (2008) propose the use of an IV under an unconditional QTE and certain identification conditions (but not under functional form assumptions). Unlike the conditional QTE with respect to *X*, the QTE is unconditional on the compliers in the Fröhlich and Melly (2008) context.

Second, Athey and Imbens (2006) have proposed the *Quantile Difference-in-Difference* (QDID). As a consequence of the inequality shown in equation 18, individual heterogeneity cannot be cancelled out with observational methods and panel data, as in a linear DID context (Khandker et al, 2010). Thus, Athey and Imbens (2006) suggest that the counterfactual distribution is equal to the difference in time of *Y* of the control group plus the pre-treatment Y of the treated group, under the debatable assumption that the counterfactual distribution over time is equal to the treatment group's; i.e.:

$$Y_0^T(\tau) + (Y_1^C(\tau) - Y_0^C(\tau))$$

Thus, Athey and Imbens compares similar individuals between groups and periods for each quantile, and then, they calculate QTE(τ).

Finally, Abrevaya and Dahl (2005) and Khandker et al (2009) propose to identify the fixed effects model under panel data with the Chamberlain (1982) model. They estimate a linear relationship between individual fixed effects and the observable characteristics, and then they estimate a pooled linear quantile regression (thus the fixed effects were eliminated in the first step).

Summing up, the quantile regression method has been used more frequently in the impact evaluation context. However, the difficulty in obtaining adequate identification strategies for its implementation with other observational methods has become a problematic barrier for its use at a widespread level.

III. Description of Programs and Variables

III.1. Introduction

This section illustrates the key characteristics of the three DIF-Puebla programs evaluated in the present investigation and their main outcome variables. This will allow, in the following section, to formulate the justification of the combined evaluation methods selected for each program.

The programs of DIF-Puebla aimed at enhancing the nutritional status of its beneficiaries, complying with the "*nutricia*"[14] quality standard, assuring community development, and fostering a correct nutrition among its beneficiaries and their families. The impact of the programs will be evaluated under four broad areas: food support, food orientation, education, and health. Though the first two areas are explicitly related to the DIF programs' main components, the other two are typically

[14] This is a high level standard of nutritional status set by the Mexican government (NOM-043-SSA2-2005). The purpose of this norm is to establish a general criterion for a proper and healthy eating habit.

analyzed outcomes in these kinds of social programs. The following sub-sections describe the programs and the variables to be evaluated.

III.2. Programs

III.2.1. DEC

The Hot School Breakfast program (DEC by its acronym in Spanish) is focused on children attending kinder, primary, secondary, and high school from public institutions of the 217 *municipios* (municipalities) of Puebla, *preferably* located in indigenous areas, rural areas, or deprived urban areas.

The beneficiaries receive a hot school breakfast every day of the schooling cycle, under "*nutricia*" standards, comprised by: 250 milliliters of skimmed milk or natural water, one hot dish of vegetables, raw cereal, legumes or meat, and at least 30 grams of fruit (fresh or dehydrated). The requisites of the program are:

- The beneficiaries should be attending a public school affiliated to the SEP (Public Education Ministry, by its acronym in Spanish).
- Their parents should create a committee that holds a constitutive act, which includes the president and vice-president names.
- Their school should be *preferably* located in a locality of high or very high marginalization degree.
- Their school should be *preferably* located in a locality where the majority of the population speaks an indigenous language.
- The beneficiary should not be receiving another nutritional program from the government.
- The school should have a physical space for installing the necessary facilities.

This program also contemplates that the beneficiary should pay a five pesos fee[15] for each meal, while the municipality should pay a maximum of 85 percent of the program expenditure.

III.2.2. DEF

The Cold School Breakfast program (DEF by its acronym in Spanish) is focused on children and teenagers attending kinder or primary school of a public school at any of the 217 municipalities of Puebla, *preferably* located in indigenous areas, rural areas or deprived urban areas.

The meal, which should comply with *nutricia* standards and should be delivered every day of the schooling cycle, comprises: 250 milliliters of semi-skimmed and ultra-pasteurized milk, 30 grams of raw cereal (oat or amaranth cookies, among others), and at least 30 grams of fruit (fresh or dehydrated). The requisites of the program are:

- The beneficiaries should be attending a public school affiliated to the SEP.
- Their parents should create a committee that holds a constitutive act, which includes the president and vice-president names.
- Their school should be *preferably* located in a locality of high or very high marginalization degree.
- Their school should be *preferably* located in a locality where the majority of the population speaks an indigenous language.
- The beneficiary should not be receiving another nutritional program from the government.

Schools are not required to ensure a physical spot to prepare and serve the cold breakfasts, thus schools with more deprived conditions may self-select into this program.

[15] Approximately 50 cents of American dollars.

Finally, this program has a recovery fee of three pesos for each meal at the beneficiary level, whereas the municipal recovery fee is the same to the DEC program.

III.2.3. INC

The "Starting a Correct Nutrition" program (INC by its acronym in Spanish) assists children between one and three years old, *preferably* located in indigenous areas, rural areas, or deprived urban areas, within the 217 municipalities of Puebla.

A monthly food package is delivered, under *nutricia* standards, comprised by: fortified milk and basic food products, such as legumes, cereals, and meat, among others. The beneficiaries should comply with the following requirements:

- To be between one and three years old.
- To be *preferably* located in a locality of high or very high marginalization degree.
- To be *preferably* located in a locality where the majority of the population speaks an indigenous language.
- Not to be receiving another nutritional program from the government.
- To comply with an economic profile applied by DIF-Puebla.

III.3. Outcome Variables by Topic and Program

III.3.1. Food Support

Food support, under *nutricia* standards, is a crucial part of the programs. The evaluation of this component is associated with the inner particularities of each program and their incidence over the beneficiaries or their households.

III.3.1.1. DEC and DEF

The effect of these programs under this topic will be analyzed through a **food insecurity index** at the household level. Martinez and Fernandez (2006) suggest that anthropometric measures are not appropriate variables to consider in impact evaluations of children attending at least primary schools since their growth indicators may reflect specific upward trends of teenagers, independently of the intervention. Considering that a great proportion of the beneficiaries of both programs are at least in the primary school (especially of DEC), we opt for the food insecurity index.

This index is created from the Food Insecurity Questionnaire of the Latin American and Caribbean Food Security Scale (FAO, 2012), which asks 15 questions about the household financial capacity to buy food; e.g. if an adult skips or reduces the breakfast, lunch or dinner size, among other questions.

This questionnaire classifies households by their food insecurity level. Each question ranges from 0 to 3, thus the aggregated index (considering the 15 questions) goes from 0 (more food security) to 45 (more food insecurity). The discrete index varies from 1 (food security) to 4 (severe food insecurity).

III.3.1.2. INC

The impact of the INC program will be evaluated by **anthropometric variables** at the beneficiary level**.** According to FAO (Latham, 2002), the main anthropometric measures used to evaluate beneficiaries in the range age of the INC are:

- **Weight for age**: this is a short-term malnutrition measure, also known as *underweight.* This is a typical variable analyzed as a result of emergency situations, such as natural disasters or economic shocks.
- **Height for age**: this is a long-term malnutrition measure, also known as *stunting*. This variable reflects the impact of repetitive infections or long-term economic changes on the accumulated nutrient ingestion over time.

- **Weight for height**: also known as *wasting*, this is malnutrition measure that combines the previous measures.

In addition, the INC program will be evaluated with the **Body Mass Index (BMI) per age**, which provides similar conclusions to the weight-for-height indicator.

These anthropometric variables are standardized by WHO 2006 child growth standards, which take well-nourished individuals from Ghana, India, Norway, Oman, Brazil, and the United States as the reference population. A z-score is obtained for each indicator, according to:

$$\text{z-score} = \frac{(\text{Observed value}) - (\text{Reference population median or mean})}{\text{Reference population standard deviation}}$$

III.3.2. Food Orientation

According to the DIF-Puebla program rules, food orientation is a crucial complement for the food supports, since it attempts to encourage a healthy life style, based on an appropriate diet and physical activity, through four approaches:

1. To develop and strengthen certain capacities and attitudes in the beneficiary's households in order to enhance her nutritional situation.
2. To identify and reinvigorate regional foods.
3. To foster an active participation of both men and women in order to create proper healthy diet habits.
4. To support household food security through ecological school farms and community canteens, in order to increase diet variety and to generate additional income sources.

Food orientation can be reflected by an adequate selection, preparation, and consumption of food in the context of an appropriate diet. Thus, food orientation will be evaluated, independently of the program, by questions regarding the household **diet** and the **habit changes** at both the level of the beneficiaries and their households.

First, food orientation will be examined by **diet diversity, variety, and quality** at the household level, based on the Healthy Food Index issued by the *Universidad Veracruzana* (2012). In particular, three indicators will be analyzed:

The first one refers to diet **diversity**, which corresponds to the inclusion of different food groups, and it is classified as:

- Diverse/complete
- Some diversity/moderated
- Non-diverse/incomplete

The second indicator refers to diet **variety**, which indicates the inclusion of different food types within the same group, and it is classified as:

- Varied
- Some Variation
- Monotonous

The third indicator agglomerates the preceding ones, obtaining the diet **quality** indicator, which is classified as:

- Complete
- Moderated
- Incomplete

Food orientation will also be evaluated by the **habit change compound index**, at both the **beneficiary** and their **household** level, through questions related to the frequency of *selection, preparation, and consumption of healthy foods*.

- **Habit changes in food selection**: these questions will attempt to capture if the orientation talks have affected the acquisition of the three groups of foods (fruits and vegetables, legumes or meat, and cereals), if these foods have been bought in the region, if ecological school farms are used, whether food is low on fat, sugar and salt or not, among others.

- **Habit changes in food preparation**: these questions will examine if the orientation talks have propitiated hygienic habits during food preparation, which cooking techniques were used, among others.

- **Habit changes in food consumption**: this part will ask about food portion sizes when eating, if each meal time is respected, if the context for eating is healthy, among others.

Finally, the **habit change compound index** is calculated, which is based on the three preceding sub-indexes.

Each indicator (i.e. **diet diversity, variety, and quality,** and the **habit change compound index, together with its sub-indexes in selection, preparation, and consumption)** will be estimated by a categorical and a continuous variable, with the purpose of obtaining more information about the causal effect of the programs.

III.3.3. Education

Nutrition and food habits of children attending school may have a direct effect on student performance. Therefore, this study will evaluate the impact on **student's marks** (only for DEC beneficiaries attending primary school), on **school absenteeism** (DEC and DEF), and on weekly hours of **extra-curricular studies** (only for DEC beneficiaries attending primary school).

III.3.4. Health

Health conditions of program beneficiaries are directly influenced by their nutritional status. Thus, this evaluation will examine the impact of the three DIF-programs on the likelihood of different diseases, spread through food, associated with the nutritional status of the treated units.

III.3.5. Summary

The following chart describes the variables to be analyzed by topic and program, as a summary of this section. It also points out whether the level of analysis is at the beneficiary or at the household level.

CHART 1: Response Variables per Program and Topic

Topic	Variable	Program	Dimension	Variable Description
Food Support	Food Insecurity Index (continuous and discrete)	DEC & DEF	Household	Comprised by 15 questions, each one ranging from 0 to 3, thus going from 0 (more food security) to 45 (more food insecurity) in aggregated terms; i.e. **the higher the index, the higher food insecurity**. The discrete index varies from 1 (food security) to 4 (severe food insecurity).
	WAZ z-score	INC	Beneficiary	Individual weight for age minus the average weight for age of the reference population, divided by the standard deviation of the reference population. A low index refers to low weight. When the index is high, it is better to observe the WHZ score. I utilize the 2006 WHO child growth standards.
	HAZ z-score	INC	Beneficiary	Individual height for age minus the average height for age of the reference population, divided by the standard deviation of the reference population. ; i.e. the higher the index, the better the child development. I utilize the 2006 WHO child growth standards.
	WHZ z-score	INC	Beneficiary	Individual weight for height minus the average weight for height of the reference population, divided by the standard deviation of the reference population. High values refers to overweight, while low values indicate emaciation. i utilize the 2006 WHO child growth standards.
	BMI for age z-score	INC	Beneficiary	The Body Mass Index is an indicator of the fat level in the body. High values indicate overweight, while low values suggest underweight.
Food Orientation	Perception of habit changes in food selection, preparation and consumption (continuous and discrete)	DEC, DEF & INC	Beneficiary & Household	I create an index based on several questions; the higher the index**, the healthier the eating behaviour**. The continuous index varies from 0 to 100, while the discrete one is a binary variable (0 or 1). At the household level, I measure i) selection; ii) preparation; iii) consumption); and iv) a weighted index on the preceding ones. At the beneficiary level, I measure i) selection; ii) consumption; and iii) a weighted index based on the preceding ones.
	Diet Diversity, Variety & Quality (continuous and discrete)	DEC, DEF & INC	Household	The eating behaviour index, utilized for evaluating the diet quality, is comprised by the measurement of two dimensions: diet diversity and diet variety. Diet diversity refers to the consumption of different food groups. Diet variety refers to the consumption of different types of food within a food group. **The higher the index, the worse the diet**. Diet diversity is measured through 7 food categories, each one valued from 0 to 10 and, in aggregated terms, ranging from 0 (diverse diet) to 70 (non-diverse diet). Its categorical variable varies from 1 (complete) to 3 (incomplete). Diet variety is measured by 6 food categories. Each survey respondent should mention 3 foods of each category (except in 2 categories, in which only 1 food should me mentioned). One unit is added for each food that is not consumed. Thus, the index varies from 0 (highest variety) to 14 (lowest variety); i.e. 4*3 + 2*1. Its categorical variable ranges from 1 (varied) to 3 (non-varied). The diet quality continuous index is the result of the addition of the diet diversity continuous index and the diet variety continuous index. Its categorical variable ranges from 1 (healthier) to 3 (less healthier).
Education	Marks	DEC	Beneficiary	Average mark in the last schooling cycle which varies from 0 to 10 (only primary school).
	School Absenteeism	DEC & DEF	Beneficiary	School Absenteeism in the last i) schooling month; and ii) schooling cycle.
	Extra-curricular studies	DEC	Beneficiary	Minutes of study outside school per week (only primary school).
Health	Diarrhea and breathing problems	DEC, DEF & INC	Beneficiary & Household	Weekly frequency of: i) diarrhea or stomach pain; and ii) breathing difficulties. **The higher the variable, the more deprived health condition**; i.e. 0 refers to non-symptoms, while 4 indicates daily-symptoms.
	Eye or gum disease or yellowish skin	DEC, DEF & INC	Beneficiary & Household	Last month frequency of: i) yellowish skin and obscured urine; ii) eyes disease symptoms; and iii) gum disease symptoms. **The higher the variable, the worse health condition**; i.e. 0 (no symptoms) and 1 (symptoms).

IV. Evaluation Methodology Choice

After having reviewed the impact evaluation methodological framework and the three DIF-programs, together with their outcome variables, I will present in this section the limitations that this research faces, and afterwards, the justification of the methodologies chosen for the impact evaluation.

IV.1. Limitations

In particular, two main limitations will be explored: i) *ex ante* versus *ex post* evaluation; and ii) the eligibility criterion[16].

Ex ante evaluations refers to those performed at the same time the program is designed; instead, *ex post* evaluations examine the programs after being designed and/or implemented. It is important to notice that the former ones are more likely to generate more accurate estimations, since: i) baseline data can be obtained; and ii) the treatment and control groups are selected before program implementation, thus more (internally and externally) valid methods can be used (e.g. randomization), under clear, transparent and difficult to manipulate selection processes (Gertler et al, 2011)[17].

The three DIF-Puebla programs analyzed in the current investigation have been designed and implemented before this analysis. The recognition of the **ex post** nature of this evaluation leads to a reduction of the array of impact evaluation methods. In particular, the experimental methods should be discarded, thus the bias in the estimations are potentially higher. Therefore, it will be used a combination of quasi-experimental methods, "based on the realities of how the program was conducted, and what data are available" (DiNardo and Lee, 2010:32). This is a

[16] Another bias that the research faces, for example, is the one generated from the fact that the direct beneficiary is not answering the questionnaire; it is rather an adult of the household.

[17] Gertler et al (2011) call *ex-ante* evaluations as "prospectives" and *ex-post*'s as "retrospectives".

common procedure when an impact evaluation is performed over: i) priority governmental programs (this is the case with the DIF-programs, in line with the *Cruzada Nacional contra el Hambre*); or ii) programs arising as a consequence of an economic crisis (Jalan and Ravaillon, 2003).

The second sizable limitation of the current investigation is the *eligibility criterion* actually followed by the DIF-Puebla authorities. It was previously stated, among the program requirements, that the beneficiaries (INC) or their schools (DEC and DEF) should be *preferably* located in localities: i) of high or very high marginalization degrees; and ii) where the majority speaks an indigenous language. These *theoretical* requirements correspond fairly well with the available data, since, for example, 85 percent of the DEC and DEF schools are located within the high and very high degree of marginalization, while 79 percent of the INC beneficiaries are found in the same degree of marginalization (Chart 2).

CHART 2: Beneficiaries per Program

	DEC		DEF		INC	
	Schools		Schools		Beneficiaries	
Marginalization	Number	%	Number	%	Number	%
Very High	59	4.0%	142	6.3%	1826	4.1%
High	1208	81.2%	1753	78.3%	33576	75.3%
Medium	147	9.9%	206	9.2%	5150	11.5%
Low	38	2.6%	108	4.8%	2083	4.7%
Very Low	36	2.4%	29	1.3%	1973	4.4%
Total	1488	100%	2238	100%	44608	100%

However, after some interviews between the UNDP-Mexico Team and the DIF-Puebla authorities, it has been unveiled that the eligibility criterion is neither strict nor exclusive in practice; rather, it follows a first-in-first-out logic due to the excess of public funds not covered by the amount of beneficiaries.

Taking into account that all school requests are accepted (if the other administrative requirements are fulfilled), these schools may have certain characteristics that systematically differ from the selected control group. For example, schools receiving the programs may have more motivated authorities and beneficiaries, and this motivation may be determining better outcomes variables, instead of the actual effect of the programs. Thus, this important evaluation limitation reveals the necessity of balancing both groups by observable characteristics, yet unobservables cannot be controlled for as a consequence of the *ex post* evaluation nature.

IV.2. Selected Methods

Due to the evaluation limitations previously mentioned, the impact evaluation will be carried out by the *propensity score* in order to balance the treatment and the control groups by observable features, and thus creating a common support for obtaining, afterwards, the casual effect. Since there is no propensity score *par excellence*[18] (as shown in the literature review), this study will use the PSM with *Stratification Matching, NN Matching* and *Kernel Matching*[19]. At the same time, the PSW will be performed with either: i) robust standard errors clustered at the locality level; and ii) block-bootstrapped standard errors, with 100 replications, also clustered at the locality level. In other words, the impact of the programs on each variable will be tested by five PS methods.

For practical reasons, as a *first condition,* I will consider that there is empirical evidence of the impact of a program on each variable when the estimated causal effect is significant (and its sign does not change) in at least three out of the five PS estimations. Second, since it is worth differentiating the confidence level of the estimations, I will create a scoring scheme; i.e. if the first condition was fulfilled, each result significant at the 90, 95 or 99 percent confidence level will receive 1.5, 1.75, or 2 points, respectively[20]. For example, if the estimation of the causal effect is significant at the 99 percent confidence level by the five PS methods, this outcome variable will have a score of 10 points. If the results are significant in two or less methods, it will be considered that there is no empirical evidence of the impact on this variable and will receive zero points (since the first condition is not complied).

[18] Except for the consensus of avoiding the PS as a covariate method.
[19] The PSM with *Radius Matching* is not presented, since several results do not converge.
[20] This is a non-linear scoring in the sense that a large premium (1.5 points) is given if the method finds the outcome variable significant at the 90 percent. Later on, if the confidence level increases, it only adds 0.25 extra points per additional block of confidence.

Finally, if a certain variable is significant at the 90 percent in two methods and at the 95 percent in a third one, it will receive 4.75 points (1.5*2 + 1.75).

Afterwards, the scores will be related to the empirical evidence found as described in Chart 3: i) no empirical evidence if the score is less than 4.5 (i.e. not even three methods provide significant coefficients at least at the 90 percent level); ii) small empirical evidence if the score is 4.5 (i.e. 3 methods at the 90 percent confidence level); iii) some empirical evidence if the score range is more than 4.5 and less than 8.75; and iv) large empirical evidence if the score is at least 8.75 (with a maximum of 10 points)[21].

CHART 3: Score for Determining the Degree of Evidence

Degree of Evidence	Range of points
Large Evidence	>=8.75
Some Evidence	> 4.5 and < 8.75
Small Evidence	4.5
No Evidence	<4.5

It is critical to point out that the **DID** method cannot be implemented on these programs, since there is not data of the outcome variables at two periods of time; thus, the casual effect will represent a simple difference between the individuals of the treatment and the control groups that lie on the common support, thereby only controlling for observables.

Additionally, the **quantile regression** will be performed over some continuous variables that are crucial for DIF authorities; i.e. student's marks in DEC and anthropometric measures in INC.

Finally, I will mention the reasons why the IV and RDD methods were not used to evaluate the programs. The difficulty in finding an appropriate instrumental variable in the context of these programs and their questionnaires leaves the **IV method** out of chances. On the one hand, as a randomized offering was not performed

[21] This methodology was created in order to summarize the large amount of results that were estimated by several PS methods. Though it is true that this methodology is subjective to the researcher point of view, it was necessary for presentation and organizational issues.

beforehand, the IV cannot be embodied by the initial random selection of eligible units. On the other hand, there were no administrative questions at the school level as a proxy for the likelihood of their students to be beneficiaries of the programs (following Arcand and Bassole, 2006).

As regards the **regression discontinuity design**, it cannot be applied due to the inexistence of a precise eligibility criterion, in practice, that may determine a clear threshold between groups. For example, if only those individuals located in a very high and high marginalization locality were selected to treatment, and the others were selected to the control group, individuals around this discontinuity could have been used for evaluating the program through RDD. However, the *first-in-first-out* logic dominates, thus this option is discarded.

V. Impact Evaluation

V.1. General Considerations

V.1.1. Standard Errors

Since PSM with Kernel Matching offers a non-parametric estimation, the standard errors may be seriously biased. The same problem arises with other parametric PSM methodologies and with the PSW, since the estimated variance from the causal effect should also include the effect: i) of the variance from the PS estimation in the first step; ii) from the creation of a subsample that fulfills the common support; and iii) of the order in which the individuals are matched when a PSM without replacement is used (Lechner, 2002; Caliendo and Kopeining, 2005; Khanker et al, 2010).

Consequently, the PSM methods will include bootstrapped standard errors, as usual. Bootstrapping takes repetitive samples from the original one, where standard errors are re-estimated in each sample, taking into account the estimations of both the PS and the structural equation. Although there is scarce evidence about how appropriate are the bootstrapped standard errors in the PSM context, this technique usually generates valid standard errors and confidence intervals (Imbens, 2004).

In particular, block-bootstrapping will be used due to the clustered structure of the variance-covariance matrix, allowing individuals within the cluster to be correlated as a result of the agglomeration (Wooldridge, 2002: 329-331), and thus avoiding biased estimations of the causal effect (Li et al, 2013).

Finally, as already mentioned, the PSW will be estimated under two different schemes: i) robust standard errors clustered at the locality level; and ii) block-bootstrapped standard errors, with 100 replications, also clustered at the locality level.

V.1.2. Control for Unobservables

The PSM balances the treatment and control groups by observables. If at least two points in time were taken, DID or a fixed effects model may be applied, thus individual heterogeneity can be controlled for. Since this data is not available for the present investigation, the results of this research may be biased by unobservables.

In order to *reduce* this source of bias, the structural estimations will contain **fixed effects at the locality level**, thus controlling for every common shock that individuals from the same locality are facing. In the case of the PS estimations, **fixed effects at the municipal level** are included. This higher aggregation level in the PS estimations was considered with the purpose of facilitating the PS estimation for each program[22].

V.1.3. PSM and PSW

Some particularities of the implementation of the PSM and the PSW will be clarified in the following paragraphs. First, a ***logit* model** will be used to determine the likelihood of participating in the program. The results by this model are pretty similar

[22] In the first place, locality fixed effects were included and the propensity scores were, in general, perfectly determined by only some localities and no other covariates. Thus, it was decided to include municipality fixed effects.

to those obtained by a probit model, though the latter has heavier tails in their distribution. In addition, these models are preferred against a linear probability model that may generate predictions out of the probability limit [0, 1] -see Smith (1997) for a discussion of the topic.

Second, following Jalan and Ravallion (2003) and Heckman, Ichimura and Todd (1997), the treatment and the control group: i) answered the **same questionnaire**; and ii) lived in the **same economic environment** in the sense that both groups are balanced by geographical terms and that there are specific estimations by only rural and only urban units. These strategies significantly increase the accuracy of the results.

Third, as suggested by Heckman, Ichimura and Todd (1997), Bryson, Dorsett and Purdon (2002), and Rubin and Thomas (1996), an **extensive list of covariates** will be used in the PS estimations, though an over-parameterized model will be avoided, in line with the literature review. Due to the *ex post* nature of the current research, the pre-treatment covariates were retrospectively obtained, thus potentially generating *recall bias*. With the purpose of addressing the problem of both the over-parameterized model and the recall bias, a simple model will be sought for the PS estimation. That is, at first, it will only include time-invariant variables. Then, it will progressively add new variables significantly correlated with the PS that, at the same time, balance the groups, as recommended by Caliendo and Kopeining (2005).

Fourth, when the PSM is used with *Kernel Matching*, a **kernel function** must be selected. This is used to weight the distance among individuals from the different groups and to perform a non-parametric *weighted least squares* estimation (Smith and Todd, 2005). The kernel function may be uniform, Epanechnikov or Gaussian, among others. This evaluation will consider a Gaussian one. In any case, this choice does not have a determinant effect on the causal effect estimation (DiNardo and Tobías, 2001).

Finally, as regards the PSW estimation, the treatment group's weight will be 1, while the control group's will be PS/(1-PS), as suggested by Hirano and Imbens (2001), Morgan and Todd (2008), and Nicholas (2008).

V.2. Covariates

In Chart 4, there is a list of ten variables included in the PS estimation for each program, while Chart 5 provides the complete list of covariates.

CHART 4: Control Variables for the Propensity Score Estimation, by Program

DEC	DEF	INC
HH age	Per Capita Food Expanditure	Marginalization Degree
Household with washing mashine	Overcrowding rate	Per Capita Food Expenditure
Household with mobile phone	Urban or rural locality	Household with refrigerator
Attend 2nd grade of Primary School	Foreing remittances received	Household with internet access
Attend 3rd grade of Primary School	Property registered for agricultural use	Belongs to the Ayotoxco de Guerrero *municipio*
Survey respondent age	Household with TV	Belongs to the Huehuetla *municipio*
Belongs to the Ayotoxco de Guerrero *municipio*	Belongs to the Cuyoaco *municipio*	Belongs to the San Nicolás de los Ranchos *municipio*
Belongs to the Chiautla *municipio*	Belongs to the Nealtican *municipio*	Belongs to the San Salvador el Seco *municipio*
Belongs to the Chignautla *municipio*		Belongs to the Tetela de Ocampo *municipio*
Belongs to the Nopalucan *municipio*		Belongs to the Zacatlán *municipio*

Note: In the DEF program, I mention the only eight variables balancing the sample. HH refers to the household head.

<u>CHART 5</u>: Control Variables

Dimension	Variable Description
Household	# of children aged 3 to 5
Household	# of household members with a disability (without including the HH)
Household	# of people older than 65
Household	% of household members working
Household	At least one household member receiving another government social program
Household	At least one household member speaks an indigenous language
Household	Drainage
Household	Dwelling deprivation (equal to 1 if dirt floor, sheet metal roof or sheet metal wall)
Household	Electric Energy
Household	Foreing remittances (equal to 1 if received)
Household	HH age
Household	HH disability (equal to 1 if having a disability)
Household	HH economic activity (equal to 1 if working)
Household	HH gender (equal to 1 if men)
Household	HH marital status (equal to 1 if having a partner)
Household	Household owner (equal to 1 if owner)
Household	Property registered for agricultural use
Household	Household with heater
Household	Household with internet
Household	Household with iron
Household	Household with mobile phone
Household	Household with refrigerator
Household	Household with TV
Household	Household with washing machine
Household	Other household member assist to the same beneficiary's shool (only used in DEC and DEF)
Household	Overcrowding Rate
Household	Per capita food expenditure
Household	Per capita income
Household	Running water
Household	Survey respondent age
Household	HH Years of schooling
Household	Years of schooling of individuals older than 14 who do not study
Beneficiary	Attend 2nd grade of Primary School (only used in DEC)
Beneficiary	Attend 3rd grade of Primary School (only used in DEC)
Beneficiary	Attend 4th grade of Primary School (only used in DEC)
Beneficiary	Attend 5th grade of Primary School (only used in DEC)
Beneficiary	Beneficiary age
Beneficiary	Beneficiary gender
Beneficiary	Minutes from house to school (only used in DEC and DEF)
Locality	Locality Fixed Effects

<u>Note</u>: HH refers to the household head.

V.3. DEC

The evaluation of the DEC program starts by comparing pre-treatment characteristics between the treatment and the control groups. The large dissimilarities between groups highlight the importance of balancing them by the PS. Chart 6 shows that individuals from the control group are situated, in 2010, in localities with a higher level of marginalization. For example, 76 percent of the control group is in a high or very high marginalized locality, while this percentage decreases to 55 percent for the treated group.

CHART 6: Marginalization Degree by Localities

Treatment Variable	Marginalization Degree per Locality in 2010					
DEC	Very Low	Low	Medium	High	Very high	Total
Control	4.34	12.43	7.62	60.49	15.12	100
Treatment	9.28	23.51	12.01	53.88	1.33	100
Total	7.54	19.61	10.47	56.2	6.18	100

Seemingly, Chart 7 illustrates that there is a higher percentage of control group units in rural than in urban areas (54 and 46 percent, respectively), as opposed to the treatment group (32 and 68 percent, respectively). This same chart shows that the percentage of people speaking an indigenous language is smaller in the treatment group (14 versus 19 percent in the control group).

CHART 7: Urban or Rural Locality and Indigenous Population

Treatment Variable	Urban or Rural Locality			At least one household member speaking an indigenous language		
DEC	Rural	Urban	Total	No	Yes	Total
Control	53.69	46.31	100	81.36	18.64	100
Treatment	31.77	68.23	100	86.13	13.87	100
Total	39.47	60.53	100	84.45	15.55	100

Finally, Chart 8 presents the pre-treatment income and food expenditure per capita averages at the level of the households. The treated units face a higher income per capita than the control group, not only by a simple average but also when survey weights are considered. As regards the per capita food expenditure, this is higher in the treatment group by a simple average, but it is slightly smaller by the weighted one (Chart 8).

CHART 8: Per Capita Income and Food Expenditure (By Household)

	Per capita income		Per capita food expenditure	
Treatment Variable: DEC	Simple Average	Weighted Average*	Simple Average	Weighted Average*
Control	611.52	611.52	348.03	348.03
Treatment	835.47	633.47	399.35	324.68

*Weighted average by survey weights. The units of the control group have a weight of 1.

In brief, these charts anticipate that the control group is more vulnerable than the treatment group. Without balancing by the PS, these differences may over-estimate the causal effect due to selection bias. Thus, the PS is estimated (Figure 6), and the remaining bias will only be generated by unobservables.

FIGURE 6: PS Estimation (DEC)

****

Algorithm to estimate the propensity score

****

The treatment is t_DEC

Variable Tratamiento DEC	Freq.	Percent	Cum.
Control	853	35.15	35.15
Tratamiento	1,574	64.85	100
Total	2,427	100	

Estimation of the propensity score

Iteration 0: log likelihood = -1573.5392
Iteration 1: log likelihood = -1422.4247
Iteration 2: log likelihood = -1418.5309
Iteration 3: log likelihood = -1418.495
Iteration 4: log likelihood = -1418.495

Logistic regression Number of obs = 2427
LR chi2(11) = 310.09
Prob > chi2 = 0.0000
Log likelihood = -1418.495 Pseudo R2 = 0.0985

t_DEC	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
edad_JH	-0.0162411	0.0054362	-2.99	0.003	-0.0268958	-0.0055864
edad_entr	-0.016585	0.0067966	-2.44	0.015	-0.0299062	-0.0032639
qk13_11_bis	0.64648	0.1036374	6.24	0	0.4433545	0.8496056
qk13_21_bis	0.231393	0.095343	2.43	0.015	0.0445243	0.4182618
yr_ed_FE1	0.6056023	0.1137068	5.33	0	0.382741	0.8284635
yr_ed_FE2	1.012068	0.1090595	9.28	0	0.7983154	1.225821
munFE5	0.7262376	0.2944506	2.47	0.014	0.1491251	1.30335
munFE7	1.089302	0.2801601	3.89	0	0.5401987	1.638406
munFE8	1.319658	0.2827063	4.67	0	0.7655642	1.873752
munFE16	0.9720547	0.2926242	3.32	0.001	0.3985219	1.545588
munFE24	-1.358603	0.3398062	-4	0	-2.024611	-0.6925949
_cons	0.8268225	0.2316838	3.57	0	0.3727306	1.280914

Note: the common support option has been selected
The region of common support is [.07811931, .96142338]

Description of the estimated propensity score in region of common support

Estimated propensity score

	Percentiles	Smallest		
1%	0.2735082	0.0781193		
5%	0.3695242	0.108476		
10%	0.4201659	0.1206623	Obs	2426
25%	0.5154398	0.138448	Sum of Wgt.	2426
50%	0.6689835		Mean	0.6487748
		Largest	Std. Dev.	0.1665424
75%	0.7787249	0.9541228		
90%	0.8559109	0.956936	Variance	0.0277364
95%	0.8977319	0.9595618	Skewness	-0.3615688
99%	0.9370616	0.9614234	Kurtosis	2.414213

****

Step 1: Identification of the optimal number of blocks

****

The final number of blocks is 8
This number of blocks ensures that the mean propensity score is not different for treated and controls in each blocks

****

Step 2: Test of balancing property of the propensity score

****

The balancing property is satisfied
This table shows the inferior bound, the number of treated and the number of controls for each block

Inferior of block of pscore	Variable Tratamiento DEC Control	Tratamiento	Total
0.0781193	9	4	13
0.2	109	63	172
0.4	340	323	663
0.6	200	343	543
0.7	124	394	518
0.8	53	204	257
0.85	9	137	146
0.9	8	106	114
Total	852	1,574	2,426

Note: the common support option has been selected

In addition, Figure 6 illustrates that the balancing test is satisfied, thus both groups are balanced by the PS; i.e. the likelihood of participation is similar in each block for the two groups. At the same time, Graph 1 illustrates the histogram of the PS for each group, thus visualizing their degree of juxtaposition and the common support area.

<u>GRAPH 1</u>: PS Histogram by Treatment Status (DEC)

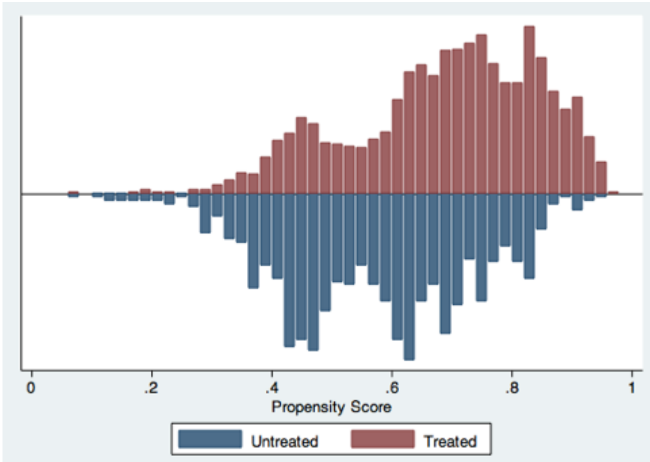
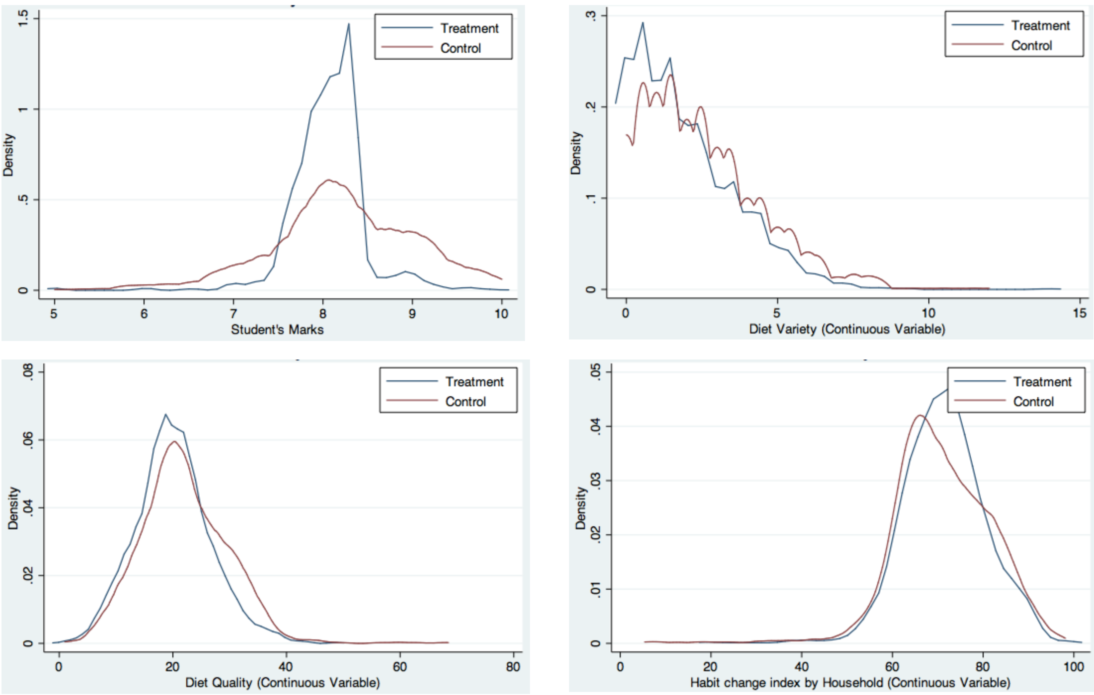


Figure 6 also illustrates that the PS estimation depends on the household head age, if the household has a washing machine, if it has a mobile phone, on the survey respondent age, on a dummy variable if attending second year of primary school, a dummy variable if attending the third year, and various municipality fixed effects, as noticed in Chart 4. In addition, it is interesting to see that only one individual was eliminated for establishing the common support area.

Figure 7 shows the density function estimations of some outcome variables through the Kernel method. The upper left graph shows that the student's marks of both groups are concentrated around the eight points and that the control group distribution is much softer than the one of the treated group. The other three illustrations from Figure 7 analyze different outcome variables from the food orientation topic. The upper right graph examines diet variety in its continuous form. The treated group is concentrated at low values of the distribution, as opposed to the control group, thus preliminary suggesting that the diet is more varied in the treatment group (0 points represent the most varied diet and 14 the least). The same occurs with the lower left graph that explores the continuous quality diet variable, which varies from 0 (more quality) to 84 (less quality). Finally, the lower right graph shows more concentration of the treated units in the higher values of the continuous index of habit changes at the beneficiary level, which ranges from 0 (worst eating habit) to 100 (best eating habit), in line with the two preceding graphs.

FIGURE 7: Kernel Density Function Estimation

(Selected Outcome Variables)



After this preliminary analysis, I will show the results of the impact evaluation of the DEC program for the overall sample. In addition, in order to capture heterogeneous effects for more specific policy implications, the causal effect will also be estimated for the following sub-samples: i) boys; ii) girls; iii) urban localities; and iv) rural localities.

Due to the great amount of results, Chart 9 only shows a summary of the significant causal effects, with the reminder that this study considers certain empirical evidence if at least three out of the five evaluation methods show a significant coefficient without changing sign. For presentation issues, this chart excludes: i) those estimations of a categorical outcome variable if also evaluated by a continuous one[23]; ii) the lower level of aggregation of the habit change compound indexes (i.e. it only includes the overall index and excludes those only referring to selection, preparation or consumption); and iii) school absenteeism in the last month, since it is less precise than the one measuring absenteeism in the last schooling cycle. In any case, all the results of the DEC program are shown at the end of this study in Annex I.

The impact of the DEC program is illustrated in Chart 9, by sample and topic. As regards the food support area, the DEC program has only a partial effect on the food insecurity index. In particular, there is a negative association between program participation and the categorical index in a range between 3 to 4 percent coming from the control group average in the general sample, thus reducing the household food insecurity perception. However, this effect is neither seen in the other sub-samples nor in the continuous index.

The program has a beneficial impact on the food orientation area, not only by different samples (general, girls, boys, urban, and rural area) but also by diverse outcome variables (household diet diversity, variety, and quality, on the one hand, and habit change perception by beneficiaries and households, on the other hand). The favorable results are more pronounced in rural areas, where the diet variety coefficient ranges from -0.39 to -0.93, equivalent to a decrease from 15 to 35 percent with respect to the weighted average of the control group in the rural sample that lies on the common support. In a gender comparison, girls are more benefited by the program. The results are significant for diet diversity, variety, and quality and for the habit change perception by beneficiaries. This last outcome variable presents the strongest evidence, since the five methods are significant at the 99 percent confidence level, thus it receives a score of ten points. The DEC program has a favorable impact on boys only through the habit change perception by beneficiaries. This impact is captured by the five methods at the 99 percent confidence level and ranges from 13 to 17 percent with respect to the control group of boys lying on the common support.

[23] The only exception is the food insecurity index in the general sample.

The impact of DEC on the education arena is quite conflictive. The program is associated with lower student's marks in the range between 2 and 3 percent (except for boys and for rural areas). Two possible interpretations may arise from this result: i) unobservable characteristics may be biasing the estimations; ii) a perverse incentive may be determining that the beneficiaries are discouraged to obtain better marks. This can happen if, for example, the beneficiaries reduce their effort in studying as a result of perceiving a long-lasting government aid. Though the first option is viable, the second one turns more likely, considering: i) the beneficial effects found in the other outcome variables; and ii) the better pre-treatment conditions found in the treated group.

Finally, in the health area, the analysis focuses on the impact on the likelihood of five diseases at both the beneficiary and their household level. The DEC is associated with an increased probability of breathing problems in boys in a range of 18-28 percent, coming from the control group weighted average. This result is also unexpected; however, there is not a great amount of evidence in this direction, since: i) only three out of the five methods suggest this result; and ii) it was neither found at the household level nor on the other samples.

Before giving an end to the DEC evaluation, and with the purpose of shedding more light on the unexpected results on student's marks, Chart 10 presents the causal effect at different points of the outcome variable distribution; i.e. on the first, second and third quartile. These **quantile regressions** will be performed through the PSW with block-bootstrapped standard errors (100 replications) under the general sample. It is important to notice that the validity of these results lies, again, on the degree of compliance of the PS assumptions.

CHART 9: Impact of the DEC Program

Program	Sample	Topic	Variable	Range	Weighted average of the control group by program and sample	Impact range						# of methods significant (min=3; max=5)	Score	Empirical Evidence of the Impact	Expected Result
						Min		Max	Min		Max				
DEC	Gral	Food Support	Food Insecurity Index by Household (*Categorical Index*)	1 (Food Security) a 4 (Severe Food Insecurity)	2.041	-0.07	to	-0.09	-3%	to	-4%	5	7.5	Some Evidence	Yes
DEC	Gral	Food Orientation	Diet quality by Household (*Continuous Index*)	0 (more healthy) to 84 (less healthy)	22.048	-0.91	to	-1.72	-4%	to	-8%	5	8.25	Some Evidence	YES
DEC	Gral	Food Orientation	Habit change perception by Beneficiary (*Continuous Index*)	0 (less healthy) to 100 (more healthy)	48.804	6.09	to	8.05	12%	to	16%	5	10	Large Evidence	YES
DEC	Gral	Food Orientation	Diet variety by Household (*Continuous Index*)	0 (more variety) to 14 (less variety)	2.379	-0.25	to	-0.5	-11%	to	-21%	5	6.5	Some Evidence	YES
DEC	Gral	Education	Student' marks in Primary School	0 to 10	8.219	-0.15	to	-0.19	-2%	to	-2%	5	9	Large Evidence	NO
DEC	Girls	Food Orientation	Diet quality by Household (*Continuous Index*)	0 (more healthy) to 84 (less healthy)	22.157	-1.47	to	-2.09	-7%	to	-9%	5	9.75	Large Evidence	YES
DEC	Girls	Food Orientation	Diet diversity by Household (*Continuous Index*)	0 (diverse diet) to 70 (non-diverse diet)	19.77	-1.04	to	-1.64	-5%	to	-8%	4	6.75	Some Evidence	YES
DEC	Girls	Food Orientation	Habit change perception by Beneficiary (*Continuous Index*)	0 (less healthy) to 100 (more healthy)	48.052	6.15	to	8.83	13%	to	18%	5	10	Large Evidence	YES
DEC	Girls	Food Orientation	Diet variety by Household (*Continuous Index*)	0 (more variety) to 14 (less variety)	2.387	-0.43	to	-0.56	-18%	to	-23%	5	8	Some Evidence	YES
DEC	Girls	Education	Student' marks in Primary School	0 to 10	8.27	-0.16	to	-0.21	-2%	to	-3%	4	6.75	Some Evidence	NO
DEC	Boys	Food Orientation	Habit change perception by Beneficiary (*Continuous Index*)	0 (less healthy) to 100 (more healthy)	49.546	6.57	to	8.49	13%	to	17%	5	10	Large Evidence	YES
DEC	Boys	Health	Beneficiary breathing difficulties (*Ordinal Categorical Variable*)	0 (never symptoms) a 4 (daily symptoms)	1.028	0.18	to	0.29	**18%**	**to**	**28%**	3	4.75	Some Evidence	NO
DEC	Rural	Food Orientation	Diet diversity by Household (*Continuous Index*)	0 (diverse diet) to 70 (non-diverse diet)	20.684	-2.9	to	-3.4	-14%	to	-16%	3	5.25	Some Evidence	YES
DEC	Rural	Food Orientation	Habit change perception by Beneficiary (*Continuous Index*)	0 (less healthy) to 100 (more healthy)	50.369	8.06	to	8.8	16%	to	17%	3	6	Some Evidence	YES
DEC	Rural	Food Orientation	Habit change perception by Household (*Continuous Index*)	0 (less healthy) to 100 (more healthy)	71.042	3.55	to	4.16	5%	to	6%	3	5.5	Some Evidence	YES
DEC	Rural	Food Orientation	Diet variety by Household (*Continuous Index*)	0 (more variety) to 14 (less variety)	2.632	-0.39	to	-0.93	-15%	to	**-35%**	4	7	Some Evidence	YES
DEC	Urban	Food Orientation	Habit change perception by Beneficiary (*Continuous Index*)	0 (less healthy) to 100 (more healthy)	47.038	7.35	to	8.61	16%	to	18%	5	10	Large Evidence	YES
DEC	Urban	Education	Student' marks in Primary School	0 to 10	8.358	-0.2	to	-0.29	-2%	to	-3%	5	10	Large Evidence	NO

CHART 10: Quantile Effects on Student' Marks (DEC)

Variable	Simple Average	Impact	Confidence Level	Impact in %
Average Marks	8.2	-0.17	***	-2.1%
1st Quartile Marks	7.8	-0.11	**	-1.4%
2nd Quartile Marks	8.2	-0.2	***	-2.4%
3rd Quartile Marks	8.9	-0.25	***	-2.8%

Note: I take the simple average of those individuals in the control group situated within the common support as the benchmark. This exercise was done over the whole DEC sample found in the common support attending primary school (N=1614). *** refers to a 99% confidence level, ** to a 95% and * to a 90%.

Chart 10 shows that the impact is negative and significant for every quartile. However, the impact is larger and more significant for the higher quartiles. In particular, the program is associated with a decrease in student's marks in 1.4 percent for the first quartile, 2.4 percent for the second one, and 2.8 for the third one. This implies that the detrimental effect did not augment initial differences.

V.4. DEF

The first step in the cold school breakfast (DEF) analysis is the comparison of the pre-treatment characteristics between groups. Chart 11 shows a similar pattern compared with the DEC program, because the control group is also more marginalized than the treated one; i.e. 62 percent of the control units are placed in localities under a high or very high level of marginalization, while this number decreases to the 50 percent in the treatment group.

CHART 11: Marginalization Degree by Localities

Treatment Variable	Marginalization Degree per Locality in 2010					
DEF	Very Low	Low	Medium	High	Very high	Total
Control	13.09	19.51	5.13	53.27	8.99	100
Treatment	11.31	27.83	11.05	47.52	2.29	100
Total	12.2	23.69	8.11	50.38	5.62	100

Chart 12 illustrates that the control group was equally balanced between urban and rural regions, while the treated units are more heavily localized in urban areas (81 percent). The same chart shows that the control group tends to speak an indigenous language with more frequency; i.e. 30 percent in the control group versus 17 percent in the treatment.

CHART 12: Urban or Rural Locality and Indigenous Population

Treatment Variable	Urban or Rural Locality			At least one household member speaking an indigenous language		
DEF	Rural	Urban	Total	No	Yes	Total
Control	49.42	50.58	100	70.09	29.91	100
Treatment	19.44	80.56	100	83.21	16.79	100
Total	34.36	65.64	100	76.68	23.32	100

Finally, Chart 13 shows that the control group has higher income and food expenditure pre-treatment levels, as opposed to the trend showed in the previous results. However, these differences between groups are not significantly different and they seem to be driven by outliers situated in very low marginalized localities, as it can be perceived in Chart 11.

CHART 13: Per Capita Income and Food Expenditure (By Household)

	Per capita income		Per capita food expenditure	
Treatment Variable: DEF	Simple Average	Weighted Average*	Simple Average	Weighted Average*
Control	791.54	791.54	413.01	413.01
Treatment	753.95	681.11	400.61	361.40

*Weighted average by survey weights. The units of the control group have a weight of 1.

In sum, it cannot be concluded that there are significant pre-treatment differences between groups. Let us take a look, then, to the PS estimation of the DEF program through Figure 8.

FIGURE 8: PS Estimation (DEF)

****

Algorithm to estimate the propensity score

****

The treatment is t_DEF

Variable Tratamiento DEF	Freq.	Percent	Cum.
Control	779	49.74	49.74
Tratamiento	787	50.26	100
Total	1,566	100	

Estimation of the propensity score

Iteration 0: log likelihood = -1082.6908

Iteration 1: log likelihood = -962.96733

Iteration 2: log likelihood = -960.78678

Iteration 3: log likelihood = -960.73683

Iteration 4: log likelihood = -960.73651

Logistic regression	Number of obs =	1562
	LR chi2(11) =	243.91
	Prob > chi2 =	0.0000
Log likelihood = -960.73651	Pseudo R2 =	0.1126

t_DEF	Coef. Std.	Err. z	P>|z|	[95%	Conf. Int	erval]
qg2_a_bis	0.8350698	0.4605008	1.81	0.07	-0.0674951	1.737635
gs_pcap	-0.000872	0.0002491	-3.5	0	-0.0013599	-0.0003834
urb	1.755296	0.1303556	13.47	0	1.499804	2.010789
hacin	0.1275488	0.0418761	3.05	0.002	0.0454732	0.2096244
qk7_bis	0.6660944	0.1391862	4.79	0	0.3932945	0.9388942
qk13_4_bis	0.6945618	0.2314994	3	0.003	0.2408313	1.148292
munFE9	2.789298	1.156879	2.41	0.016	0.5218572	5.056739
munFE15	-1.405447	0.3907245	-3.6	0	-2.171253	-0.6396407
_cons	-3.638344	0.9654596	-3.77	0	-5.53061	-1.746078

Note: the common support option has been selected

The region of common support is [.11391711, .98868107]

Description of the estimated propensity score in region of common support

Estimated propensity score

	Percentiles	Smallest		
1%	0.1393415	0.1139171		
5%	0.2030709	0.1157953		
10%	0.2295938	0.1169754	Obs	1556
25%	0.3144983	0.1236284	Sum of Wgt.	1556
50%	0.5508411		Mean	0.5028477
		Largest	Std. Dev.	0.1906026
75%	0.6450998	0.9724343		
90%	0.7312543	0.985017	Variance	0.0363294
95%	0.7794431	0.987877	Skewness	-0.2120626
99%	0.8359768	0.9886811	Kurtosis	1.980531

****

Step 1: Identification of the optimal number of blocks

****

The final number of blocks is 5

This number of blocks ensures that the mean propensity score is not different for treated and controls in each blocks

****

Step 2: Test of balancing property of the propensity score

****

The balancing property is satisfied

This table shows the inferior bound, the number of treated and the number of controls for each block

Inferior of block of pscore	Variable Tratamiento DEF		
	Control	Tratamiento	Total
0.1139171	59	15	74
0.2	312	123	435
0.4	225	238	463
0.6	169	360	529
0.8	8	47	55
Total	773	783	1,556

Note: the common support option has been selected

This figure shows that ten individuals are discarded from the original sample (four from the treatment and six from the control group) to balance both groups in terms of observables and to find the common support illustrated in Graph 2. Figure 8, as well as Chart 4, shows that DEF participation is explained by foreign remittances, per capita food expenditure, a dummy variable equal to 1 if the locality belongs to an urban area or 0 otherwise, the overcrowding rate, if the household has a TV, if the property is registered for agricultural use, and two municipality fixed effects.

Figure 9 shows the estimations of: i) the histograms of two categorical outcome variables; and ii) the Kernel Epanechnikov density function of two continuous outcome variables. The upper graphs show the diarrhea or stomach pain weekly frequency for households and beneficiaries (left and right chart, respectively), measured through an ordinal categorical variable. In both cases, a higher proportion of treated units has less symptoms. In the lower charts, from left to right, the density functions of the habit change perception variable by households and beneficiaries, respectively, are deployed. The left chart shows that a larger proportion of households of the treated group is located in the upper part of the distribution (i.e. better eating habits), while the right chart does not infer substantial differences at the beneficiary level.

GRAPH 2: PS Histogram by Treatment Status (DEF)

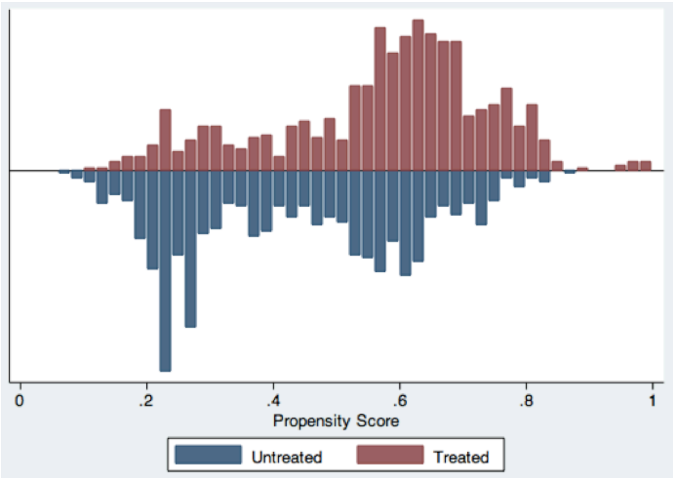


FIGURE 9: Preliminary Graphic Analysis

(Selected Outcome Variables)

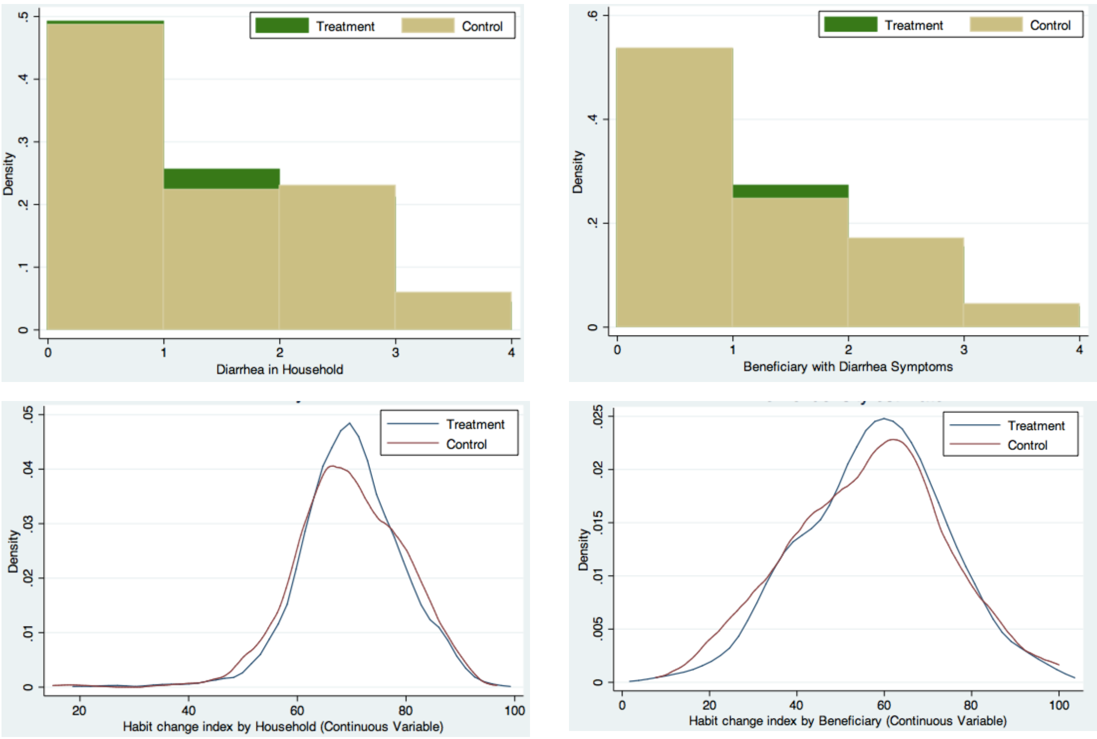


Chart 14 describes the impact of the DEF program. As it was done with the DEC analysis, the main results can be visualized in this chart, while the overall results may be found in Annex II.

First, there is no significant association between DEF participation and the food insecurity index (under the food support topic). The categorical index is inversely related to program participation, as expected, but only in the rural sample in two out of the five evaluation methods (Annex II).

Second, DEF is associated with better food orientation outcomes at the *household* level, measured by the habit change perception and diet diversity variables, not only for the general sample (3-4 percent) but also for girls (1-10 percent), and urban areas (4-5 percent). However, this program is associated with worse food orientation outcomes at the *beneficiary* level, measured by habit change perception, except in rural areas where no significant effects were found.

Third, in the education field[24], DEF is associated with an increase in school absenteeism between 42 to 45 percent in the general sample; yet, no significant results were found in the sub-samples. This result is in line with the detrimental effects of DEC on education. Presumably the same potential interpretations can be provided: i) results are biased by unobservables; or ii) there may be perverse incentives of the program on their beneficiaries. Though it was highlighted that the second option may be more viable for DEC, it is not necessarily the same in this program, considering that this result was significant in three out of the five evaluation methods for only the general sample.

Finally, as regards the health area, DEF is associated with lower diarrhea symptoms in *girls*, not only for the beneficiaries (34-62 percent) but also for their households (29-60 percent). Having found this effect at both levels, the impact of this outcome variable for girls is reinforced. On the other hand, this effect was not found in the other samples.

[24] Student's marks are not evaluated in the DEF program, since a large amount of beneficiaries (from DIF reports) were attending kinder school.

CHART 14: Impact of the DEF Program

Program	Sample	Topic	Variable	Range	Weighted average of the control group by program and sample	Impact range						# of methods significant (min=3; max=5)	Score	Empirical Evidence of the Impact	Expected Result
						Min		Max	Min		Max				
DEF	Gral	Food Orientation	Habit change perception by Household (*Continuous Index*)	0 (less healthy) to 100 (more healthy)	70.467	2.31	to	3.06	**3%**	**to**	**4%**	5	7.5	Some Evidence	YES
DEF	Gral	Education	School Absenteeism in last schooling cycle	0 to 87 days	2.118	0.9	to	0.95	42%	to	45%	3	6	Some Evidence	NO
DEF	Gral	Food Orientation	Habit change perception by Beneficiary (*Continuous Index*)	0 (less healthy) to 100 (more healthy)	62.054	-6.26	to	-8.97	-10%	to	-14%	5	10	Large evidence	NO
DEF	Girls	Food Orientation	Diet diversity by Household (*Continuous Index*)	0 (diverse diet) to 70 (non-diverse diet)	18.57	-0.17	to	-1.89	**-1%**	**to**	**-10%**	4	7.5	Some Evidence	YES
DEF	Girls	Food Orientation	Habit change perception by Household (*Continuous Index*)	0 (less healthy) to 100 (more healthy)	69.941	4.07	to	5.92	6%	to	8%	5	9.5	Large evidence	YES
DEF	Girls	Health	Beneficiary's Diarrhea symptoms (*Ordinal Categorical Variable*)	0 (never symptoms) a 4 (daily symptoms)	0.742	-0.25	to	-0.46	-34%	to	-62%	3	5.75	Some Evidence	YES
DEF	Girls	Health	Diarrhea symptoms in the Household (*Ordinal Categorical Variable*)	0 (never symptoms) a 4 (daily symptoms)	0.712	-0.21	to	-0.43	-29%	to	-60%	5	9.25	Large evidence	YES
DEF	Girls	Food Orientation	Habit change perception by Beneficiary (*Continuous Index*)	0 (less healthy) to 100 (more healthy)	60.855	-8.35	to	-8.6	-14%	to	-14%	3	6	Some Evidence	NO
DEF	Boys	Food Orientation	Habit change perception by Beneficiary (*Continuous Index*)	0 (less healthy) to 100 (more healthy)	63.231	-8.35	to	-10.4	-13%	to	-16%	5	10	Large evidence	NO
DEF	Urban	Food Orientation	Habit change perception by Household (*Continuous Index*)	0 (less healthy) to 100 (more healthy)	69.148	2.43	to	3.24	**4%**	**to**	**5%**	5	8.25	Some Evidence	YES
DEF	Urban	Food Orientation	Habit change perception by Beneficiary (*Continuous Index*)	0 (less healthy) to 100 (more healthy)	60.5	-6.57	to	-9.22	-11%	to	-15%	5	9.75	Large evidence	NO

V.5. INC

The evaluation of the "Starting a Correct Nutrition" (INC) program also begins by comparing pre-treatment differences between groups. As occurred with the previous programs, Chart 15 shows that the control group has a larger proportion of individuals residing in localities with high or very high marginalization levels (66 versus 53 percent in the treatment group).

CHART 15: Marginalization Degree by Localities

Treatment Variable	Marginalization Degree per Locality in 2010					
INC	Very Low	Low	Medium	High	Very high	Total
Control	19.18	7.35	7.47	62.89	3.11	100
Treatment	26.58	7.35	13.25	50.6	2.22	100
Total	23.57	7.35	10.9	55.6	2.58	100

Chart 16 indicates that a larger proportion of the treatment group is located in urban areas -almost ten percentage points higher than the control group. In addition, this chart shows that the control group presents a higher proportion of individuals speaking an indigenous language than the treatment group (21 versus 15 percent, respectively).

CHART 16: Urban or Rural Locality and indigenous Population

Treatment Variable	Urban or Rural Locality			At least one household member speaking an indigenous language		
INC	Rural	Urban	Total	No	Yes	Total
Control	31.51	68.49	100	79.33	20.67	100
Treatment	23.59	76.41	100	84.7	15.3	100
Total	26.81	73.19	100	82.51	17.49	100

Finally, Chart 17 shows that the treatment group presents higher incomes and food expenditures under the different types of analysis. These results are in line with the previous pre-treatment comparisons, suggesting that the control group is more vulnerable than the treatment group. This highlights the importance of balancing the groups through the propensity score estimation, which is presented in Figure 10.

CHART 17: Per Capita Income and Food Expenditure

(By Household)

	Per capita income		Per capita food expenditure	
Treatment Variable: INC	Simple Average	Weighted Average*	Simple Average	Weighted Average*
Control	639.05	601.48	337.73	320.81
Treatment	732.47	653.81	412.94	363.37

*Weighted average by survey weights.

FIGURE 10: PS Score Estimation (INC)

****

Algorithm to estimate the propensity score

****

The treatment is t_INC

Variable Tratamiento INC	Freq.	Percent	Cum.
Control	803	40.7	40.7
Tratamiento	1,170	59.3	100
Total	1,973	100	

Estimation of the propensity score

Iteration 0: log likelihood = -1332.724

Iteration 1: log likelihood = -1247.2492

Iteration 2: log likelihood = -1244.8238

Iteration 3: log likelihood = -1244.7819

Iteration 4: log likelihood = -1244.7819

Logistic regression Number of obs = 1972

LR chi2(11) = 175.88

Prob > chi2 = 0.0000

Log likelihood = -1244.7819 Pseudo R2 = 0.0660

t_INC	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
gr_marg	-0.904285	0.044841	-2.02	0.044	-0.1783152	-0.0025417
gs_pcap	0.0009947	0.0002326	4.28	0	0.0005387	0.0014507
qk13_9_bis	0.2611674	0.1033049	2.53	0.011	0.0586936	0.4636413
qk13_17_bis	-0.817624	0.3854174	-2.12	0.034	-1.573028	-0.0622194
munFE5	1.217157	0.4589025	2.65	0.008	0.3177246	2.116589
munFE10	-2.177375	0.6121422	-3.56	0	-3.377152	-0.9775981
munFE21	1.795524	0.4148813	4.33	0	0.9823715	2.608676
munFE23	1.221084	0.3805662	3.21	0.001	0.4751882	1.96698
munFE30	1.548392	0.3176818	4.87	0	0.9257468	2.171037
munFE37	2.10694	0.6140735	3.43	0.001	0.9033785	3.310502
_cons	0.053992	0.2079435	0.26	0.795	-0.3535697	0.4615536

Note: the common support option has been selected

The region of common support is [.09095891, .9475187]

Description of the estimated propensity score in region of common support

Estimated propensity score

	Percentiles	Smallest		
1%	0.407748	0.0909589		
5%	0.4503615	0.0930361		
10%	0.465625	0.095091	Obs	1955
25%	0.5018167	0.0958719	Sum of Wgt.	1955
50%	0.5595568		Mean	0.5972042
		Largest	Std. Dev.	0.1333654
75%	0.6730652	0.9384003		
90%	0.8171089	0.9423464	Variance	0.0177863
95%	0.8579912	0.9423464	Skewness	0.3906595
99%	0.9043326	0.9475187	Kurtosis	3.708735

****

Step 1: Identification of the optimal number of blocks

****

The final number of blocks is 6

This number of blocks ensures that the mean propensity score is not different for treated and controls in each blocks

****

Step 2: Test of balancing property of the propensity score

****

The balancing property is satisfied

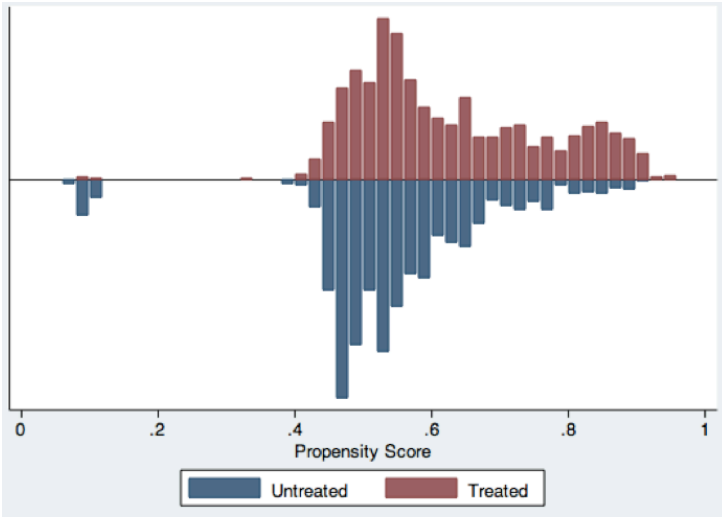
This table shows the inferior bound, the number of treated and the number of controls for each block

Inferior of block of pscore	Variable Tratamiento INC		
	Control	Tratamiento	Total
0.0909589	12	3	15
0.2	2	1	3
0.4	262	204	466
0.5	299	415	714
0.6	181	351	532
0.8	30	195	225
Total	786	1,169	1,955

Note: the common support option has been selected

The common support condition reduces the sample in 18 individuals (1 treated and 17 from the control group) to the total amount of 1955. Program participation is estimated by the marginalization degree of the locality, the per capita food expenditure, if the household has a refrigerator, if it has internet access, plus several municipality fixed effects. The degree of juxtaposition is illustrated in Graph 3, which shows, for instance, a small number of units of both groups with low levels of the PS.

<u>GRAPH 3</u>: PS Score Histogram by Treatment Status (INC)



<u>FIGURE 11</u>: Kernel Density Function Estimation

(Selected Outcome Variables)

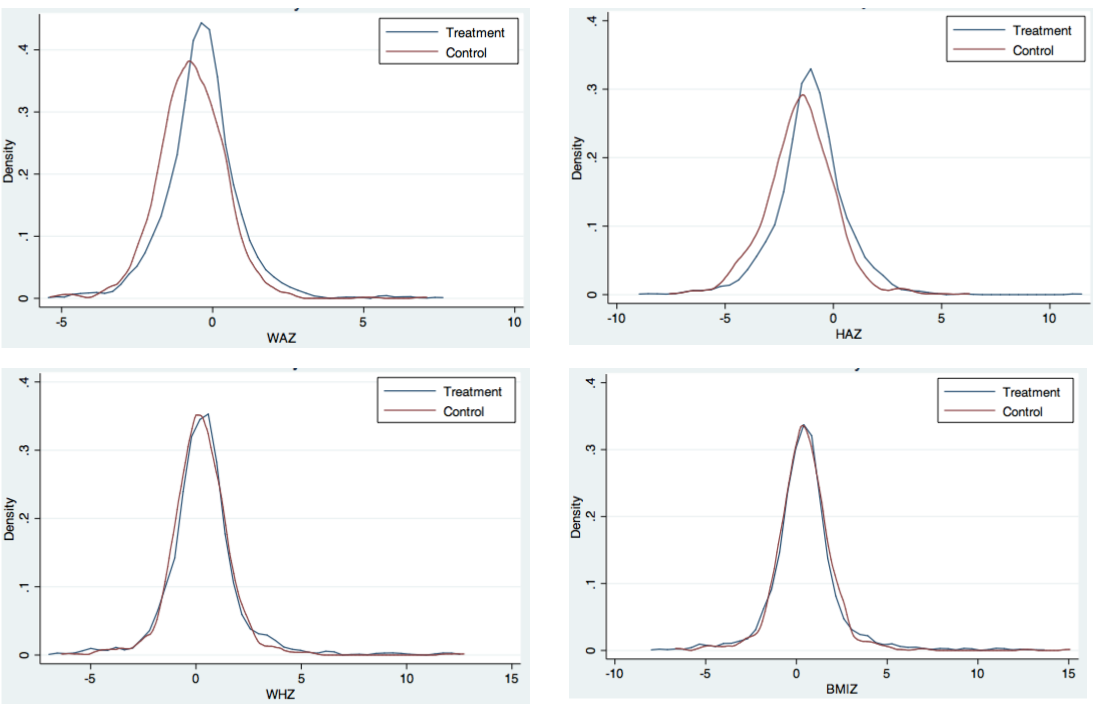


Figure 11 examines the four anthropometric outcome variables analyzed in this evaluation. As expected, their averages are located around zero since they are standardized with respect to the reference population. In addition, all the variables present certain bias to the right, in the sense that there are some outliers in the right tails of the distributions. The upper graphs suggest that the treatment group has larger weight-for-age and height-for-age z-scores, presumably indicating promising results of the program at this regard. However, the lower graphs do not suggest substantial differences between groups in the weight-for-height z-score and the BMI per age z-score.

Chart 18 contains the main effects of the INC impact evaluation, while the whole results are presented in Annex III. **The INC has a positive impact on the anthropometric measures**, reflecting beneficial effects of the <u>food supports</u> on the beneficiaries. Specifically, the participation in the program is associated with higher height-for-age- z-scores or HAZ (i.e. 24-31 percent in the general sample, 26-37 percent for girls, 16-25 percent for boys, and 33-35 percent in urban areas), except for those beneficiaries in rural areas, where the results were insignificant. These results determine that the beneficiaries get closer to the international reference population average, thus leaving behind the "very short" threshold. This can also be appreciated in Chart 19, which shows: i) where is located the average z-score of each group for each variable in blue (e.g. HAZ-T refers to the HAZ average of the treatment group)[25]; and ii) the significant variables shaded in grey (e.g. HAZ averages from the rural sample were not shaded, since they were insignificant). In the first column, where HAZ is presented, it can be seen that not a single group from any sample is located in the short stature range (i.e. HAZ<-2). At the same time, this column indicates that the program generates a jump of range in the general sample (from the control group average between -2 and -1 to the treatment average between -2 to 1), which is more pronounced for girls and urban areas (from -2 to -1 to -1 to 1). Boys receive a positive impact of the program but this is not translated into a jump of range.

[25] The control group average consists in the z-score weighted average of those individuals located in the common support. The treatment group average is the control group average plus the range of the INC impact.

<u>CHART 18</u>: Impact of the INC Program

Program	Sample	Topic	Variable	Range	Weighted average of the control group by program and sample	Impact range						# of methods significant (min=3; max=5)	Score	Empirical Evidence of the Impact	Expected Result
						Min		Max	Min		Max				
INC	Gral	Food Support	HAZ	-8.7 to 11.2	-1.431	0.35	to	0.44	24%	to	31%	4	7.75	Some Evidence	YES
INC	Gral	Food Support	WAZ	-5.38 to 7.43	-0.671	0.2	to	0.29	30%	to	43%	5	10	Large evidence	YES
INC	Gral	Health	Beneficiary's Gum disease symptoms (*Binary Categorical Variable*)	0 (without symptoms) to 1 (with symptoms)	0.033	-0.001	to	-0.002	**-3%**	**to**	**-6%**	3	4.5	Small Evidence	YES
INC	Girls	Food Support	HAZ	-8.7 to 11.2	-1.303	0.34	to	0.48	26%	to	37%	4	7.5	Some Evidence	YES
INC	Girls	Food Support	WAZ	-5.38 to 7.43	-0.616	0.3	to	0.39	49%	to	63%	5	10	Large evidence	YES
INC	Girls	Food Orientation	Habit change perception by Beneficiary (*Continuous Index*)	0 (less healthy) to 100 (more healthy)	54.409	2.38	to	3.64	**4%**	**to**	**7%**	5	8.25	Some Evidence	YES
INC	Girls	Health	Beneficiary's Gum disease symptoms (*Binary Categorical Variable*)	0 (without symptoms) to 1 (with symptoms)	0.039	-0.03	to	-0.04	**-77%**	**to**	**-103%**	3	6	Some Evidence	YES
INC	Girls	Health	Yellowish skin in the Household (*Binary Categorical variable*)	0 (without symptoms) to 1 (with symptoms)	0.035	0.03	to	0.04	**86%**	**to**	**114%**	5	8.5	Some Evidence	NO
INC	Boys	Food Support	HAZ	-8.7 to 11.2	-1.549	0.25	to	0.39	16%	to	25%	5	9	Large evidence	YES
INC	Boys	Food Support	WAZ	-5.38 to 7.43	-0.721	0.2	to	0.23	28%	to	32%	4	7.5	Some Evidence	YES
INC	Rural	Food Support	BMI	-7.74 to 15.07	0.353	0.49	to	0.57	139%	to	161%	3	4.5	Small Evidence	YES
INC	Rural	Food Support	WAZ	-5.38 to 7.43	-0.735	0.37	to	0.56	50%	to	76%	5	9.75	Large evidence	YES
INC	Rural	Food Support	WHZ	-6.74 to 12.73	0.141	0.42	to	0.56	298%	to	397%	5	8	Some Evidence	YES
INC	Rural	Food Orientation	Diet quality by Household (*Continuous Index*)	0 (more healthy) to 84 (less healthy)	22.208	-1.39	to	-2.21	-6%	to	-10%	4	7	Some Evidence	YES
INC	Rural	Food Orientation	Habit change perception by Beneficiary (*Continuous Index*)	0 (less healthy) to 100 (more healthy)	54.957	3.03	to	4.51	**6%**	**to**	**8%**	4	7	Some Evidence	YES
INC	Rural	Food Orientation	Diet variety by Household (*Continuous Index*)	0 (more variety) to 14 (less variety)	2.502	-0.08	to	-0.82	**-3%**	**to**	**-33%**	5	10	Large evidence	YES
INC	Rural	Health	Beneficiary's Gum disease symptoms (*Binary Categorical Variable*)	0 (without symptoms) to 1 (with symptoms)	0.046	-0.06	to	-0.07	**-130%**	**to**	**-152%**	3	4.75	Some Evidence	YES
INC	Urban	Food Support	HAZ	-8.7 to 11.2	-1.347	0.44	to	0.47	33%	to	35%	4	7.75	Some Evidence	YES
INC	Urban	Food Support	WAZ	-5.38 to 7.43	-0.608	0.21	to	0.23	35%	to	38%	4	6.25	Some Evidence	YES

Chart 18 also shows that the INC has a positive impact on the weight-for-age z-scores or WAZ (i.e. 30-43 percent in the general sample, 49-63 percent for girls, 28-32 percent for boys, 50-76 percent in rural areas, and 35-38 percent urban areas). Though these results get the beneficiaries closer to the reference population, these improvements are not enough to produce a range jump for any sample (Chart 19, second column).

The INC has a positive effect on the weight-for-height z-score or WHZ (300-400 percent) and the BMI per age z-score (140-160 percent) only in rural areas. Even though the control group average is higher than zero, as opposed to the other anthropometric variables (Chart 18), these increases do not suggest likely overweight or obesity problems (Chart 19).

As regards food orientation, Chart 18 indicates that the INC has a beneficial effect on girls (an increase in the habit change perception variable from 4 to 7 percent) and on rural areas (6-8 percent increase in the habit change perception variable for beneficiaries, and a decrease of the diet quality and variety variables in the range of 6-10 and 3-33 percent, respectively). However, there are no significant effects in the other samples.

Lastly, in the health area, program participation is associated with lower gum disease symptoms in the beneficiary, not only in the general sample (3-6 percent) but also for girls (77-103 percent) and rural areas (130-152 percent). By contrary, INC is associated with higher yellowish skin symptoms in households in the sample of girls (86-114 percent). Considering that this last effect was only found in households (not in the beneficiaries) in one out of the five samples, this may be generated by unobservables not captured by the PS estimation.

CHART 19: Z-Score Indicators in INC

Z-score	HAZ	WAZ	WHZ	BMI by Age	Sample
> 3	Very tall	Likely overweight but better evaluated by WHZ	Obesity	Obesity	General
> 2			Overweight	Overweight	
> 1			Likely Overweight	Likely Overweight	
0	HAZ-T	WAZ-C // WAZ-T	WHZ-C // WHZ-T	IMC-C // IMC-T	
< -1	HAZ-C				
< -2	Short (stunting)	Underweight	Wasted	Wasted	
< -3	Very Short (severe	Severe underweight	Severe wasted	Severe wasted	
> 3	Very tall	Likely overweight but better evaluated by WHZ	Obesity	Obesity	Boys
> 2			Overweight	Overweight	
> 1			Likely Overweight	Likely Overweight	
0		WAZ-C // WAZ-T	WHZ-C // WHZ-T	IMC-C // IMC-T	
< -1	HAZ-C // HAZ-T				
< -2	Short	Underweight	Wasted	Wasted	
< -3	Very Short	Severe underweight	Severe wasted	Severe wasted	
> 3	Very tall	Likely overweight but better evaluated by WHZ	Obesity	Obesity	Girls
> 2			Overweight	Overweight	
> 1			Likely Overweight	Likely Overweight	
0	HAZ-T	WAZ-C // WAZ-T	WHZ-C // WHZ-T	IMC-C // IMC-T	
< -1	HAZ-C				
< -2	Short	Underweight	Wasted	Wasted	
< -3	Very Short	Severe underweight	Severe wasted	Severe wasted	
> 3	Very tall	Likely overweight but better evaluated by WHZ	Obesity	Obesity	Urban Localities
> 2			Overweight	Overweight	
> 1			Likely Overweight	Likely Overweight	
0	HAZ-T	WAZ-C // WAZ-T	WHZ-C // WHZ-T	IMC-C // IMC-T	
< -1	HAZ-C				
< -2	Short	Underweight	Wasted	Wasted	
< -3	Very Short	Severe underweight	Severe wasted	Severe wasted	
> 3	Very tall	Likely overweight but better evaluated by WHZ	Obesity	Obesity	Rural Localities
> 2			Overweight	Overweight	
> 1			Likely Overweight	Likely Overweight	
0		WAZ-C // WAZ-T	WHZ-C // WHZ-T	IMC-C // IMC-T	
< -1	HAZ-C // HAZ-T				
< -2	Short	Underweight	Wasted	Wasted	
< -3	Very Short	Severe underweight	Severe wasted	Severe wasted	

Note: the average z-score of each group for each variable are in blue (e.g. HAZ-T refers to the HAZ average of the treatment group). The significant variables are shaded in grey (e.g. both HAZ-C and HAZ-T from the rural sample were not shaded, since they were insignificant).

As a final step, the impact evaluation of the INC program contemplates quantile regressions on anthropometric measures for the general sample. As performed in the DEC evaluation, the heterogeneous effects will be evaluated for the first, second, and third quartile through the PSW with bootstrapped standard errors (100 replications).

CHART 20: Quantile Impact on Anthropometric Measures (INC)

Variable	Impact	Confidence Level
Panel A: HAZ		
HAZ average	0.44	***
HAZ 1st Q	0.48	***
HAZ 2nd Q	0.36	***
HAZ 3th Q	---	---
Panel B: WAZ		
WAZ average	0.29	***
WAZ 1st Q	0.27	**
WAZ 2nd Q	0.29	***
WAZ 3th Q	0.34	***
Panel C: WHZ		
WHZ average	0.06	
WHZ 1st Q	---	---
WHZ 2st Q	0.05	
WHZ 3st Q	0.04	
Panel D: BMIZ		
BMIZ average	-0.01	
BMIZ 1st Q	-0.02	
BMIZ 2nd Q	0.01	
BMIZ 3th Q	-0.06	

Note: This exercise was performed over the total units located in the common support region. *** refers to a 99% confidence level, ** to a 95% and * to a 90%.

Panel A, Chart 20, shows the differential impact on HAZ for the first and second quartile (results on the third quartile are not provided since the estimations do not converge for that point of the distribution). The causal effect for the first quartile is higher than for the second quartile (0.48 versus 0.36 respectively), which implies an extra benefit to those with worst initial measures. By contrary, Panel B suggests that the higher the quartiles, the larger the impact of INC on WAZ. Though the opposite would be desirable, these effects are not leading to obesity problems for the higher quartiles, since all of them depart from lower values with respect to the reference population. Finally, Panel C and D do not find significant results on WHZ and BMI per age z-score.

VI. Final Remarks

This document presents the impact evaluation of three nutritional programs of DIF-Puebla: Hot School Breakfast (DEC), Cold School Breakfast (DEF), and Starting a Correct Nutrition (INC). For this purpose, it examines, first, the main impact evaluation methods available in the literature. Based on this analysis, and on the particular characteristics of the programs, the most appropriate evaluation methods are proposed.

By five variations of the *Propensity Score Matching* and *Weighting*, the programs are evaluated under five samples: i) general sample; ii) boys; iii) girls; iv) urban localities; and v) rural localities. Taking into account the great amount of possible results, the outcome variables found significant in at least three out of the five methods are considered as providing empirical evidence of the impact. In addition, a scoring scheme is devised in order to determine: i) non-empirical evidence; ii) small evidence; iii) some evidence; and iv) large evidence.

In brief, DEC has: i) a beneficial impact on food orientation outcomes at the beneficiary and their household levels throughout different samples and estimations; ii) a marginal favorable effect on food security by households; iii) a detrimental effect on student's marks under different samples, which is larger for higher quartiles; and iv) a negative effect on breathing disease symptoms for boys (though there is not large empirical evidence about this result, since only three out of the five methods determine this result in only the boys sample).

DEF presents: i) a promising impact on food orientation outcomes on households, but unfavorable for their beneficiaries; ii) non-significant effects on food security; iii) a deleterious effect on school absenteeism on the general sample, but no effect on the sub-samples; and iv) a reduction in diarrhea symptoms in girls, not only at the beneficiary but also at their household level.

Finally, the INC generates: i) a beneficial impact on growth indicators (specifically on height-for-age and weight-for-age z-scores), consistent throughout different samples (except for rural areas) and quartiles, and with more intensity on girls; ii) a favorable effect on food orientation outcomes for girls and for rural areas (beneficiaries and households); and iii) a reduction of gum disease symptoms for the beneficiaries in three samples (the general one, girls and rural areas), though higher yellowish skin symptoms for households in the general sample.

This evaluation determines strong *policy implications.* On the one hand, it adds substantial empirical evidence of the beneficial effects of nutritional programs on growth indicators. In addition, it provides some evidence about the favorable impact of this kind of programs on food orientation outcomes, such as eating habit changes or diet diversity, variety, and quality variables. On the other hand, it unveils only marginal effects on food security and detrimental effects on the educational arena (specifically on student's marks). Finally, it does not postulate conclusive impacts on health.

This impact evaluation also provides useful *recommendations* for the DIF-policy makers. In the DEC and DEF programs, it is recommended to get deeper into the benefits of education and disease prevention within the food orientation talks. At the same time, it is proposed to revise the size and quality of the food support, since it was found small evidence about the beneficial effect on food security in the DEC program and no evidence in the DEF program. Finally, as regards the DEF program, it is also recommended to improve the food orientation talks, specifically in urban areas, focused on eating habit changes and better diets.

As regards the INC, it has proved to present sizeable beneficial effects on their beneficiaries and households. However, specific attention should be placed into rural areas, since their beneficiaries have not presented higher HAZ and WAZ measures, while the impact on WHZ and the BMI per age was significant, which eventually may lead to overweight problems. At the same time, it is important to focus on those children with initially worse growth conditions, considering the heterogeneous effects found at distinct points of the outcome variable distributions. Lastly, as suggested for the previous programs, it is recommended to improve the food orientation talks with the purpose of preventing diseases, improving eating habits, and enhancing diet diversity, variety, and quality.

VII. References

Abadie, A. Angrist, J.D. and Imbens, G.W. 2002. Instrumental variables estimates of the effect of subsidized training on the quantiles of trainee earnings, *Econometrica*, 70, 91-117.

Abrevaya, J. and Dahl, C.M. 2008. The Effects of Birth Inputs on Birthweight: Evidence from Quantile Estimation on Panel Data. *Journal of Business and Economic Statistics*, 26 (4): 379–97.

Angrist, J.D. and Imbens, G.W. 1994. Identification and Estimation of local Average Treatment Effects. *Econometrica*, 62(2):467-475.

Arcand, J.L. and Bassole, L. 2006. Does Community Driven Development Work? Evidence from Senegal. CERDI, CNRS.

Athey, S. and Imbens, G.W. 2006. Identification and Inference in Nonlinear Difference-in- Differences Models. *Econometrica*, 74 (2): 431–97.

Austin, P.C. 2011. An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. *Multivariate Behavioral Research*, 46:399-424.

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

Buchinsky, M. 1998. Recent Advances in Quantile Regression Models: A Practical Guide for Empirical Research. *Journal of Human Resources*, 33(1): 88–126.

Buddelmeyer, H. and Skoufias, E. 2004. *An Evaluation of the Performance of Regression Discontinuity Design on PROGRESA*. World Bank Policy Research Working Paper 3386, IZA Discussion Paper 827, World Bank, Washington, DC.

Caliendo, M. and Kopeinig, S. 2005. *Some Practical Guidance for the Implementation of Propensity Score Matching*. IZA Discussion Papers 1588, Institute for the Study of Labor (IZA).

Chamberlain, G. 1982. Multivariate Regression Models for Panel Data. *Journal of Econometrics*, 18 (1): 5–46.

Cook, T.D., Shadish, W.R. and Wong, V.C. 2006. *Within Study Comparisons of Experiments and Non-Experiments: Can they help decide on Evaluation Policy*. Mimeo, Northwestern University.

DiNardo, J. and Lee, D.S. 2011. *Program Evaluation and Research Designs*. In Handbook of Labor Economics, Chapter 5, Volume 4, Part A, pp. 463-536.

DiNardo, J. and Tobias, J. 2001. Nonparametric Density and Regression Estimation. *Journal of Economic Perspectives*, 15(4), 11–28.

Dobash, R.P., Dobash, R.E., Cavanagh, K., and Lewis, R. 1999. A Research Evaluation of British Programmes for Violent Men. *Journal of Social Policy,* 28(2): 205-233.

Duflo, E., Glennerster, R. and Kremer, M. 2011. *Using Randomization in Development Economics Research: A Toolkit*. Discussion Paper 6059, Centre for Economic Policy Research, UK.

FAO 2012. Latin American and Caribbean Food Security Scale. The Food Insecurity Questionnaire. Romme, Italy.

Gertler, P. J., Martinez, S., Premand, P., Rawlings, L.B. and Vermeersch, C.M.J. 2011. *Impact Evalution in Practice*. Washington DC: World Bank.

Glazerman, S., Levy D. and Myers D. 2003. *Nonexperimental Replications of Social Experiments: A Systematic Review*. Princeton, NJ: Mathematica Policy Research, Inc.

Glewwe, P. and Jacoby, H.G. 1995. An Economic Analysis of Delayed Primary School Enrollment in a Low Income Country: The Role of Early Childhood Nutrition. *Review of Economic Statistics,* 77(1):156–69.

Heckman, J. J., Ichimura, H. and Todd, P. 1997. Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme. *Review of Economic Studies* 64 (4): 605–54.

________, LaLonde R. and Smith, J. 1999. *The Economics and Econometrics of Active Labor Market Programs*. In Handbook of Labor Economics, vol. 3, ed. Orley Ashenfelter and David Card, 1865–2097. Amsterdam: North-Holland.

________, Smith, J. and Clements, N. 1997. Making the most out of programme evaluations and social experiments: Accounting for heterogeneity in programme impacts, *Review of Economic Studies,* 64, 487-535.

Holland, P. W. 1986. Statistics and Causal Inference. *Journal of the American Statistical Association*, 81:945-960.

Hirano, K. and Imbens, G. 2001, Estimation of Causal Effects Using Propensity Score Weighting: An Application of Data on Right Hear Catherization. *Health Services anf Outcomes Research Methodology,* 2, 259-278.

Imbens, G. 2004. Nonparametric Estimation of Average Treatment Effects under Exogeneity: A review.National Bureau of Economic Research, Technical Working Paper 294.

Imbens, G. and Angrist, J. 1994. Identification and estimation of local average treatment effects. *Econometrica* 62, 467-476.

_________, and Lemieux T. 2008. The regression discontinuity design: Theory and applications. *Journal of Econometrics,* 144 (2).

Khandker, S.R., Bakht, Z. and Koolwal, G.B. 2009. The Poverty Impact of Rural Roads: Evidence from Bangladesh. *Economic Development and Cultural Change*, 57 (4): 685–722.

Khandker, S.R., Koolwal, G.B. and Samad H.A. 2010. *Handbook on Impact Evaluation*. Washington DC: World Bank.

Koenker, R and Bassett G.Jr. 1978. Regression Quantiles. *Econometrika*, 46(1):33-50.

Lalonde, R.J. 1986. Evaluating the Econometric Evaluations of Training Programs Using Experimental Data. *American Economic Review*, 76(4): 602–620.

Latham, M. 2002. *Nutrición Humana en el mundo en desarrollo*. Organización para la Agricultura y la Alimentación FAO, Colección FAO: Alimentación y nutrición N° 29.

Lechner, M. 2002. Some practical issues in the evaluation of heterogenous labour market programmes by matching methods. *Journal of the Royal Statistical Society*, 165, 59–82.

Lee, D. 2008. Randomized Experiments from Non-random Selection in U.S. House Elections. *Journal of Econometrics*, 142(2):675-697.

_______ and Lemieux. 2010. Regression Discontinuity Designs in Economics, *Journal of Economic Literature*, American Economic Association, 48(2):281-355.

Li F, Zaslavsky, A.M. and Landrum, M.B. 2013. Propensity Score Weighting with Multilevel Data. Duke University and Harvard Medical School. Unpublished Document.

Lunceford, J.K. and Davidian, M. 2004. Stratification and weighting via the propensity score in estimation of causal treatment effects: a comparative study. *Statistics in Medicine*, 23: 2937–2960.

Martínez, R. and Fernández, A. 2006. *Modelo de Análisis del Impacto Social y Económico de la Desnutrición Infantil en América Latina*. Serie Manuales 52, CEPAL, Santiago de Chile.

Morgan, S. L. and Todd, J. L. 2008. A diagnostic routine for the detection of consequential heterogeneity of causal effects. *Sociological Methodology*, 38, 231–281.

Nichols A. 2008. Erratum and discussion of propensity-score reweighting. *The Stata Journal*, 8, Number 4: 532–539.

Olken, B. 2005. *Monitoring Corruption: Evidence from a Field Experiment in Indonesia*. NBER Working Paper No. 11753, National Bureau of Economic Research.

Ravallion, M. 2008. Evaluating Anti-Poverty Programs. En Handbook of Development Economics, vol. 4, ed. T. Paul Schultz y John Strauss, 3787–846. Amsterdam: North-Holland.

Robins, J.M., Rotnitzky, A. and Zhao, L.P. 1995. Analysis of semiparametric regression-models for repeated outcomes in the presence of missing data. *Journal of the American Statistical Association*, 90: 106–121.

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

Rosenbaum, P.R. 2002. *Observational Studies*. New York and Berlin: Springer-Verlag.

Rubin, D.B. and Thomas, N. 1996. Matching Using Estimated Propensity Scores: Relating Theory to Practice, *Biometrics*, 52, 249–264.

Rubin, D.B. 2001. Using propensity scores to help design observational studies: Application to the tobacco litigation. *Health Services and Outcomes Research Methodology*, 2, 169–188.

Rubin, D.B. 2004. On principles for modeling propensity scores in medical research. *Pharmaco-epidemiology Drug Safety*, 13, 855–857.

Sefton, T., Byford, S., McDaid, D., Hills, J. and Knapp, M. 2002. *Making the Most of It: Economic Evaluation in the Social Welfare Field*. Joseph Rowntree Foundation, York.

Smith, H. 1997. Matching with Multiple Controls to Estimate Treatment Effects in Observational Studies. *Sociological Methodology*, 27, 325–353.

Smith, J. and Todd, P. 2005. Does Matching Overcome LaLonde's Critique of Nonexperimental Estimators? *Journal of Econometrics*, 125 (1–2): 305–53.

Todd, P. 2007. *Evaluating Social Programs with Endogenous Program Placement and Selection of the Treated*. En Handbook of Development Economics, vol. 4, ed. T. Paul Schultz y John Strauss, 3847–94. Amsterdam: North-Holland.

Universidad Veracruzana. 2012. "Validación y Estandarización de la Herramienta para Focalizar adecuadamente la Entrega o Asignación de los programas de la Estrategia integral de Asistencia Social Alimentaria (EIASA)", Documento Rector, Facultad de Nutrición-Xalapa, Junio 2012, México.

Wooldridge, J.M. 2002. *Econometric Analysis of Cross-Section and Panel Data*. MIT Press, Cambridge, MA.

VIII. Annex

VIII.1. DEC Results by Sample

DEC General

	Food Support						Food Orientation																				
Methodology	Food insecurity by Household (*Continuous index*)			Food insecurity by Household (*Categorical index*)			Habit change perception on food selection by Household (Continuous Index)			Habit change perception on food selection by Household (Categorical Index)			Habit change perception on food preparation by Household (Continuous Index)			Habit change perception on food preparation by Household (Categorical Index)			Habit change perception on food consumption by Household (Continuous Index)			Habit change perception on food consumption by Household (Categorical Index)			Habit change perception by Household (Continuous Index)		
	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N
PSM Kernel	-1.09	*	2426	**-0.09**	*	**2426**	-0.13	*	2426	-0.04		2426	0.704		2426	0.03		2426	0.778		2426	**0.074**	**	**2426**	0.545		2426
PSM Nearest Neighbor	-1.03		2176	**-0.09**	*	**2176**	-0.25		2175	-0.04		2175	0.621		2179	0.017		2179	-0.46		2170	**0.027**		**2170**	0.062		2166
PSM Stratification	-0.96		2426	**-0.08**	*	**2426**	-0.67		2426	-0.06		2426	0.684		2426	0.03		2426	0.779		2426	**0.073**		**2426**	0.388		2426
PSW robust & cluster s.e.	-0.62		2386	**-0.07**	*	**2386**	1.557		2375	0.025		2375	0.455		2393	0.021		2393	0.014		2369	**0.06**	*	**2369**	0.579		2351
PSW Bootstrapped s.e.	-0.62		2386	**-0.07**	*	**2386**	1.557		2375	0.025		2375	0.455		2393	0.021		2393	0.014		2369	**0.06**	*	**2369**	0.579		2351

	Food Orientation																										
Methodology	Habit change perception by Household (Categorical Index)			Habit change perception on food selection by Beneficiary (*Continuous Index*)			Habit change perception on food selection by Beneficiary (*Categorical Index*)			Habit change perception on food consumption by Beneficiary (*Continuous Index*)			Habit change perception on food consumption by Beneficiary (*Categorical Index*)			Habit change perception by Beneficiary (*Continuous Index*)			Habit change perception by Beneficiary (*Categorical Index*)			Diet Diversity by Household (*Continuous Index*)			Diet Diversity by Household (*Continuous Index*)		
	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N
PSM Kernel	0.075	*	2426	**7.708**	***	**2426**	**0.13**	***	**2426**	**7.14**	***	**2426**	**0.155**	***	**2426**	**8.154**	***	**2426**	**0.041**	***	**2426**	-1.21	*	2426	-0.06	*	2426
PSM Nearest Neighbor	0.065		2166	**8.04**	***	**2176**	**0.15**	***	**2176**	**5.887**	**	**2168**	**0.132**	**	**2168**	**7.845**	***	**2165**	**0.189**	***	**2165**	-1.2		2179	-0.06		2179
PSM Stratification	0.069		2426	**8.022**	***	**2426**	**0.138**	***	**2426**	**7.254**	***	**2426**	**0.16**	***	**2426**	**8.503**	***	**2426**	**0.186**	***	**2426**	-1.22		2426	-0.06		2426
PSW robust & cluster s.e.	0.093		2351	**7.173**	***	**2389**	**0.133**	***	**2389**	**3.038**		**2357**	**0.07**		**2357**	**6.087**	***	**2353**	**0.126**	***	**2353**	-0.66		2390	-0.03		2393
PSW Bootstrapped s.e.	0.093	**	2351	**7.174**	***	**2389**	**0.133**	***	**2389**	**3.038**		**2357**	**0.07**		**2357**	**6.087**	***	**2353**	**0.126**	***	**2353**	-0.66		2390	-0.03		2393

	Food Orientation												Education												Health		
Methodology	Diet Variety by Household (*Continuous Index*)			Diet Variety by Household (*Continuous Index*)			Diet Quality by Household (*Continuous Index*)			Diet Quality by Household (*Continuous Index*)			Student' Marks in Primary School			School Absence in last month (kinder & primary school)			School Absence in last schooling cycle (kinder & primary school)			Extra-curricular studies			Diarrhea symptoms in the Household (Ordinal Categorical Variable)		
	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N
PSM Kernel	**-0.5**	**	**2426**	**-0.1**	**	**2426**	**-1.72**	**	**2426**	0.00		2426	**-0.15**	*	**1626**	**0.404**	***	**2426**	0.257		2426	7.3		1626	0.088		2426
PSM Nearest Neighbor	**-0.49**	*	**2179**	**-0.12**	**	**2179**	**-1.69**	**	**2179**	0.00		2179	**-0.19**	***	**1304**	**0.409**	***	**2175**	0.352		2179	-15.8		1300	0.024		2179
PSM Stratification	**-0.48**	**	**2426**	**-0.11**	**	**2426**	**-1.7**	**	**2426**	0.00		2426	**-0.15**	*	**1626**	**0.427**	***	**2426**	0.294		2426	7.61		1626	0.114		2426
PSW robust & cluster s.e.	**-0.25**	*	**2393**	**-0.04**		**2393**	**-0.91**	*	**2390**	0.00		2393	**-0.17**	***	**1614**	**0.223**		**2384**	-0.07		2391	29.4		1606	-0.08		2392
PSW Bootstrapped s.e.	**-0.25**	*	**2393**	**-0.04**		**2393**	**-0.91**	*	**2390**	0.00		2393	**-0.17**	***	**1614**	**0.223**		**2384**	-0.07		2391	29.4	*	1606	-0.08		2392

	Health																										
Methodology	Breathing difficulties in the Household *(Ordinal Categorical Variable*)			Yellowish skin in the Household (*Binary Categorical variable*)			Eyes disease symptoms in the Household (Binary Categorical *Variable*)			Gum disease symptoms in the Household (Binary Categorical *Variable*)			Beneficiary's Diarrhea symptoms (Ordinal Categorical Variable)			Beneficiary's Breathing difficulties (Ordinal Categorical Variable)			Beneficiary's Yellowish skin (Binary Categorical variable)			Beneficiary's Eyes disease symptoms (Binary Categorical Variable)			Beneficiary's Gum disease symptoms (Binary Categorical Variable)		
	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N
PSM Kernel	0.166	*	2426	0.006		2426	0.009		2426	0.011		2426	0.102		2426	0.174	*	2426	-0.01		2426	0.003		2426	-0.01		2426
PSM Nearest Neighbor	0.154		2179	0.004		2179	0.021		2179	0.021		2179	0.05		2179	0.16		2179	-0.01		2179	0.008		2179	0.004		2179
PSM Stratification	0.195	*	2426	0.005		2426	0.017		2426	0.015		2426	0.127		2426	0.191	*	2426	-0.02		2426	0.005		2426	-0.01		2426
PSW robust & cluster s.e.	-0.05		2390	0.03		2391	0.04		2393	0.017		2393	-0.04		2392	-0.03		2389	0.022		2393	0.034	*	2393	0.006		2393
PSW Bootstrapped s.e.	-0.05		2390	0.03		2391	0.04	*	2393	0.018		2393	-0.04		2392	-0.03		2389	0.022		2393	0.034	*	2393	0.006		2393

Note: * is significant at the 90% confidence level, ** is significant at the 95% and *** is significant at the 99%. Estimations with bootstrapped standard errors are replicated 100 times. Robust standard errors are clustered at the locality level. The causal effect is a simple difference, since the response variables are only captured in a single point of time. The propensity score estimations include municipality fixed effects, while the structural equations consider locality fixed effects.

DEC Boys

	Food Support						Food Orientation																				
Methodology	Food insecurity by Household (*Continuous index*)			Food insecurity by Household (*Categorical index*)			Habit change perception on food selection by Household (Continuous Index)			Habit change perception on food selection by Household (Categorical Index)			Habit change perception on food preparation by Household (Continuous Index)			Habit change perception on food preparation by Household (Categorical Index)			Habit change perception on food consumption by Household (Continuous Index)			Habit change perception on food consumption by Household (Categorical Index)			Habit change perception by Household (Continuous Index)		
	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N
PSM Kernel	-0.9		1212	-0.08		1212	0.253		1212	-0.04		1212	1.457	*	1212	**0.042**	*	**1212**	1.297		1212	0.069	*	1212	1.275		1212
PSM Nearest Neighbor	-1.12		1089	-0.08		1089	0.006		1090	-0.05		1090	1.258		1090	**0.05**		**1090**	1.809		1084	0.079		1084	1.311		1084
PSM Stratification	-0.67		1212	-0.61		1212	-0.3		1212	-0.06		1212	1.074		1212	**0.051**		**1212**	1.895		1212	0.1	*	1212	1.095		1212
PSW robust & cluster s.e.	-0.22		1190	-0.04		1190	2.04		1191	0.066		1191	1.493		1194	**0.061**	*	**1194**	0.398		1179	0.067		1179	1.366		1176
PSW Bootstrapped s.e.	-0.22		2386	-0.04		2386	2.04		1191	0.066		1191	1.493		1194	**0.061**	*	**1194**	0.398		2386	0.067		2386	1.366	*	1176

	Food Orientation																										
Methodology	Habit change perception by Household (Categorical Index)			Habit change perception on food selection by Beneficiary (*Continuous Index*)			Habit change perception on food selection by Beneficiary (*Categorical Index*)			Habit change perception on food consumption by Beneficiary (*Continuous Index*)			Habit change perception on food consumption by Beneficiary (*Categorical Index*)			Habit change perception by Beneficiary (*Continuous Index*)			Habit change perception by Beneficiary (*Categorical Index*)			Diet Diversity by Household (*Continuous Index*)			Diet Diversity by Household (*Continuous Index*)		
	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N
PSM Kernel	**0.097**	**	**1212**	**8.52**	***	**1212**	**0.169**	***	**1212**	5.776	**	1212	**0.142**	**	**1212**	**8.041**	***	**1212**	**0.202**	***	**1212**	-0.79		1212	-0.02		1212
PSM Nearest Neighbor	**0.094**	*	**1084**	**9.217**	***	**1090**	**0.199**	***	**1090**	4.63		1084	**0.12**	*	**1084**	**8.174**	***	**1084**	**0.236**	***	**1084**	-0.16	*	1090	-0.07		1090
PSM Stratification	**0.086**		**1212**	**8.192**	***	**1212**	**0.177**	***	**1212**	3.032	*	1212	**0.143**	**	**1212**	**8.489**	***	**1212**	**0.214**	***	**1212**	-0.87		1212	-0.03		1212
PSW robust & cluster s.e.	**0.099**	**	**1176**	**8.142**	***	**1193**	**0.162**	***	**1193**	2.86		1175	**0.067**		**1175**	**6.574**	***	**1174**	**0.162**	***	**1174**	-0.35		1191	0.019		1194
PSW Bootstrapped s.e.	**0.099**	**	**1176**	**8.142**	***	**1193**	**0.162**	***	**1193**	2.86		2386	**0.067**		**2386**	**6.574**	***	**1174**	**0.162**	***	**1174**	-0.35		1191	0.019		1194

	Food Orientation												Education												Health		
Methodology	Diet Variety by Household (*Continuous Index*)			Diet Variety by Household (*Continuous Index*)			Diet Quality by Household (*Continuous Index*)			Diet Quality by Household (*Continuous Index*)			Student' Marks in Primary School			School Absence in last month (kinder & primary school)			School Absence in last schooling cycle (kinder & primary school)			Extra-curricular studies			Diarrhea symptoms in the Household (Ordinal Categorical Variable)		
	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N
PSM Kernel	-0.54	*	1212	-0.09		1212	-1.34		1212	0.00		1212	-0.13		807	**0.478**	***	**1212**	0.256		1212	17.94		807	0.126		1212
PSM Nearest Neighbor	-0.46	*	1090	-0.1		1090	-2.01	**	1090	0.00		1090	-0.06		643	**0.456**	***	**1087**	0.312		1090	15.56		643	0.101		1090
PSM Stratification	-0.44		1212	-0.09		1212	-1.32		1212	0.00		1212	-0.12		807	**0.511**	***	**1212**	0.245		1212	18.35		807	0.135		1212
PSW robust & cluster s.e.	-0.07		1194	0.02		1194	-0.42		1191	0.01		1194	-0.17	**	807	**0.478**	*	**1191**	-0.01		1194	17.88		802	-0.04		1194
PSW Bootstrapped s.e.	-0.07		2386	0.02		2386	-0.42		1191	0.01		1194	-0.17	**	807	**0.478**	*	**1191**	-0.01		1194	17.88		802	-0.04		1194

	Health																										
Methodology	Breathing difficulties in the Household *(Ordinal Categorical Variable)*			Yellowish skin in the Household (*Binary Categorical variable*)			Eyes disease symptoms in the Household (Binary Categorical *Variable*)			Gum disease symptoms in the Household (Binary Categorical *Variable*)			Beneficiary's Diarrhea symptoms (Ordinal Categorical Variable)			Beneficiary's Breathing difficulties (Ordinal Categorical Variable)			Beneficiary's Yellowish skin (Binary Categorical variable)			Beneficiary's Eyes disease symptoms (Binary Categorical Variable)			Beneficiary's Gum disease symptoms (Binary Categorical Variable)		
	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N
PSM Kernel	0.182	*	1212	0.018		1212	0.028		1212	0.025		1212	0.129		1212	**0.177**	*	**1212**	-0.01		1212	0.001		1212	-0.01		1212
PSM Nearest Neighbor	0.18		1090	0.027		1090	0.035	*	1090	0.028		1090	0.184		1090	**0.285**	**	**1090**	-0.02		1090	0.008		1090	0.007		1090
PSM Stratification	0.197	*	1212	0.022		1212	0.033		1212	0.027		1212	0.154		1212	**0.212**	*	**1212**	-0.03		1212	0.005		1212	-0		1212
PSW robust & cluster s.e.	-0.1		1192	0.081	***	1194	0.051		1194	0.042	*	1194	-0.02		1194	**-0.09**		**1192**	0.041		1194	0.029		1194	0.006		1194
PSW Bootstrapped s.e.	-0.1		1192	0.081	***	1194	0.051		1194	0.042	*	1194	-0.03		1194	**-0.09**		**1192**	0.041		1194	0.029	*	1194	0.006		1194

Note: * is significant at the 90% confidence level, ** is significant at the 95% and *** is significant at the 99%. Estimations with bootstrapped standard errors are replicated 100 times. Robust standard errors are clustered at the locality level. The causal effect is a simple difference, since the response variables are only captured in a single point of time. The propensity score estimations include municipality fixed effects, while the structural equations consider locality fixed effects.

DEC Girls

	Food Support						Food Orientation																				
Methodology	Food insecurity by Household (*Continuous index*)			Food insecurity by Household (*Categorical index*)			Habit change perception on food selection by Household (Continuous Index)			Habit change perception on food selection by Household (Categorical Index)			Habit change perception on food preparation by Household (Continuous Index)			Habit change perception on food preparation by Household (Categorical Index)			Habit change perception on food consumption by Household (Continuous Index)			Habit change perception on food consumption by Household (Categorical Index)			Habit change perception by Household (Continuous Index)		
	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N
PSM Kernel	-1.27	**	1212	-0.1	*	1212	-0.54		1212	-0.05		1212	-0.12		1212	0.017		1212	0.233		1212	0.079	**	1212	-0.24		1212
PSM Nearest Neighbor	-1.04		1061	-0.11	*	1061	-1.08		1056	-0.04		1056	0.022		1063	0.012		1063	0.062		1059	0.039		1059	-0.35		1052
PSM Stratification	-1.2	*	1212	-0.09		1212	-1.13		1212	-0.07		1212	0.206		1212	0.011		1212	-0.02		1212	0.062	**	1212	-0.31		1212
PSW robust & cluster s.e.	-0.97		1196	-0.1		1196	1.173		1184	-0.03		1184	-0.87		1199	-0.02		1199	-0.29		1190	0.054		1190	-0.33		1175
PSW Bootstrapped s.e.	-0.97		1196	-0.1		1196	1.173		1184	-0.03		1184	-0.87		1199	-0.02		1199	-0.29		1190	0.051		1190	-0.33		1175

	Food Orientation																										
Methodology	Habit change perception by Household (Categorical Index)			Habit change perception on food selection by Beneficiary (*Continuous Index*)			Habit change perception on food selection by Beneficiary (*Categorical Index*)			Habit change perception on food consumption by Beneficiary (*Continuous Index*)			Habit change perception on food consumption by Beneficiary (*Categorical Index*)			Habit change perception by Beneficiary (*Continuous Index*)			Habit change perception by Beneficiary (*Categorical Index*)			Diet Diversity by Household (*Continuous Index*)			Diet Diversity by Household (*Continuous Index*)		
	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N
PSM Kernel	0.053		1212	**6.898**	***	**1212**	**0.092**	**	**1212**	**8.547**	***	**1212**	**0.169**	***	**1212**	**8.282**	***	**1212**	**0.164**	***	**1212**	**-1.64**	***	**1212**	**-0.11**	***	**1212**
PSM Nearest Neighbor	0.037		1052	**6.307**	***	**1061**	**0.089**	*	**1061**	**8.338**	***	**1055**	**0.177**	***	**1055**	**7.768**	***	**1053**	**0.159**	***	**1053**	**-1.23**		**1063**	**-0.07**		**1063**
PSM Stratification	0.049		1212	**7.15**	***	**1212**	**0.101**	**	**1212**	**9.17**	***	**1212**	**0.192**	***	**1212**	**8.832**	***	**1212**	**0.167**	***	**1212**	**-1.59**	**	**1212**	**-0.1**	***	**1212**
PSW robust & cluster s.e.	0.091		1175	**6.732**	***	**1196**	**0.114**	**	**1196**	**3.706**		**1182**	**0.082**		**1182**	**6.15**	***	**1179**	**0.108**	*	**1179**	**-1.04**	*	**1199**	**-0.09**	**	**1199**
PSW Bootstrapped s.e.	0.091		1175	**6.732**	***	**1196**	**0.114**	**	**1196**	**3.038**		**1182**	**0.07**		**1182**	**6.15**	***	**1179**	**0.108**	*	**1179**	**-1.04**	*	**1199**	**-0.09**	**	**1199**

	Food Orientation												Education												Health		
Methodology	Diet Variety by Household (*Continuous Index*)			Diet Variety by Household (*Continuous Index*)			Diet Quality by Household (*Continuous Index*)			Diet Quality by Household (*Continuous Index*)			Student' Marks in Primary School			School Absence in last month (kinder & primary school)			School Absence in last schooling cycle (kinder & primary school)			Extra-curricular studies			Diarrhea symptoms in the Household (Ordinal Categorical Variable)		
	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N
PSM Kernel	**-0.45**	*	**1212**	**-0.11**	***	**1212**	**-2.09**	***	**1212**	-0.01	*	1212	**-0.17**	**	**816**	0.327	**	1212	0.257		1212	-2.48		816	0.05		1212
PSM Nearest Neighbor	**-0.56**	**	**1063**	**-0.14**	***	**1063**	**-1.79**	**	**1063**	-0.01		1063	**-0.13**		**641**	0.31		1062	0.37		1063	0.718		641	0.008		1063
PSM Stratification	**-0.5**	*	**1212**	**-0.12**	***	**1212**	**-2.08**	***	**1212**	-0.01	*	1212	**-0.16**	*	**816**	0.336	**	1212	0.311		1212	-2.67		816	0.091		1212
PSW robust & cluster s.e.	**-0.43**	**	**1199**	**-0.1**	***	**1199**	**-1.47**	***	**1199**	-0.01		1199	**-0.21**	**	**807**	-0.04		1193	-0.14		1197	33.7		804	-0.09		1198
PSW Bootstrapped s.e.	**-0.43**	**	**1199**	**-0.1**	***	**1199**	**-1.47**	***	**1199**	-0.01		1199	**-0.21**	**	**807**	-0.04		1193	-0.14		1197	33.7		804	-0.09		1198

	Health																										
Methodology	Breathing difficulties in the Household *(Ordinal Categorical Variable)*			Yellowish skin in the Household (*Binary Categorical variable*)			Eyes disease symptoms in the Household (Binary Categorical *Variable*)			Gum disease symptoms in the Household (Binary Categorical *Variable*)			Beneficiary's Diarrhea symptoms (Ordinal Categorical Variable)			Beneficiary's Breathing difficulties (Ordinal Categorical Variable)			Beneficiary's Yellowish skin (Binary Categorical variable)			Beneficiary's Eyes disease symptoms (Binary Categorical Variable)			Beneficiary's Gum disease symptoms (Binary Categorical Variable)		
	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N
PSM Kernel	0.15		1212	-0.01		1212	-0.1		1212	-0		1212	0.073		1212	0.171		1212	-0.01		1212	0.005		1212	-0		1212
PSM Nearest Neighbor	0.123		1063	-0.02		1063	0.012		1063	0.009		1063	-0.06		1063	0.193		1063	0.007		1063	0.018		1063	-0.01		1063
PSM Stratification	0.199		1212	-0.01		1212	0		1212	0.001		1212	0.098		1212	0.187		1212	-0.01		1212	0.006		1212	-0.01		1212
PSW robust & cluster s.e.	-0		1198	-0.02		1197	0.019		1199	-0.01		1199	-0.01		1198	0.047		1197	-0		1199	0.034	**	1199	0.003		1199
PSW Bootstrapped s.e.	-0		1198	-0.02		1197	0.019		1199	-0.01		1199	-0.01		1198	0.047		1197	-0		1199	0.034	**	1199	0.003		1199

Note: * is significant at the 90% confidence level, ** is significant at the 95% and *** is significant at the 99%. Estimations with bootstrapped standard errors are replicated 100 times. Robust standard errors are clustered at the locality level. The causal effect is a simple difference, since the response variables are only captured in a single point of time. The propensity score estimations include municipality fixed effects, while the structural equations consider locality fixed effects.

DEC Urban

Methodology	Food Support						Food Orientation																				
	Food insecurity by Household (*Continuous index*)			Food insecurity by Household (*Categorical index*)			Habit change perception on food selection by Household (Continuous Index)			Habit change perception on food selection by Household (Categorical Index)			Habit change perception on food preparation by Household (Continuous Index)			Habit change perception on food preparation by Household (Categorical Index)			Habit change perception on food consumption by Household (Continuous Index)			Habit change perception on food consumption by Household (Categorical Index)			Habit change perception by Household (Continuous Index)		
	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N
PSM Kernel	-0.5		1467	-0.03		1467	-2.18		1467	-0.08	*	1467	-0.51		1467	0.034		1467	0.235		1467	0.063	*	1467	-0.91		1467
PSM Nearest Neighbor	-1.13		1381	-0.07		1381	-2.5		1379	-0.08		1379	-0.3		1383	0.007		1383	-0.33		1378	0.03		1378	-1.1		1374
PSM Stratification	-0.54		1467	-0.03		1467	-2.4		1467	-0.09	*	1467	-0.37		1467	0.035		1467	0.377		1467	0.067	*	1467	-0.88		1467
PSW robust & cluster s.e.	-0.09		1437	-0.03		1437	1.489		1427	0.033		1427	-0.31		1442	0.002		1442	0.385		1434	0.065	*	1434	0.374		1419
PSW Bootstrapped s.e.	-0.09		1437	-0.03		1437	1.489		1427	0.033		1427	-0.31		1442	0.002		1442	0.385		1434	0.065	*	1434	0.374		1419

Methodology	Food Orientation																										
	Habit change perception by Household (Categorical Index)			Habit change perception on food selection by Beneficiary (*Continuous Index*)			Habit change perception on food selection by Beneficiary (*Categorical Index*)			Habit change perception on food consumption by Beneficiary (*Continuous Index*)			Habit change perception on food consumption by Beneficiary (*Categorical Index*)			Habit change perception by Beneficiary (*Continuous Index*)			Habit change perception by Beneficiary (*Categorical Index*)			Diet Diversity by Household (*Continuous Index*)			Diet Diversity by Household (*Continuous Index*)		
	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N
PSM Kernel	0.035		1467	**7.173**	***	**1467**	**0.097**	**	**1467**	**8.13**	**	**1467**	**0.188**	***	**1467**	**8.264**	***	**1467**	**0.162**	***	**1467**	0.192		1467	-0		1467
PSM Nearest Neighbor	0.046		1374	**7.165**	***	**1383**	**0.105**	**	**1383**	**6.317**	**	**1378**	**0.17**	**	**1378**	**7.512**	***	**1378**	**0.157**	***	**1378**	0.429		1383	0.026		1383
PSM Stratification	0.032		1467	**7.36**	***	**1467**	**0.102**	**	**1467**	**8.577**	**	**1467**	**0.205**	***	**1467**	**8.614**	***	**1467**	**0.172**	***	**1467**	0.189		1467	-0		1467
PSW robust & cluster s.e.	0.082		1419	**7.95**	***	**1441**	**0.15**	***	**1441**	**4.84**	*	**1427**	**0.086**		**1427**	**7.35**	***	**1426**	**0.144**	***	**1426**	-0.68		1439	-0.02		1442
PSW Bootstrapped s.e.	0.082		1419	**7.95**	***	**1441**	**0.133**	***	**1441**	**4.84**		**1427**	**0.086**		**1427**	**7.35**	***	**1426**	**0.144**	***	**1426**	-0.68	*	1439	-0.02		1442

Methodology	Food Orientation												Education												Health		
	Diet Variety by Household (*Continuous Index*)			Diet Variety by Household (*Continuous Index*)			Diet Quality by Household (*Continuous Index*)			Diet Quality by Household (*Continuous Index*)			Student' Marks in Primary School			School Absence in last month (kinder & primary school)			School Absence in last schooling cycle (kinder & primary school)			Extra-curricular studies			Diarrhea symptoms in the Household (Ordinal Categorical Variable)		
	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N
PSM Kernel	-0.22		1467	-0.07		1467	-0.03		1467	**0.01**	*	**1467**	**-0.29**	***	**924**	**0.382**	**	**1467**	-0.07		1467	13.32		924	0.171		1467
PSM Nearest Neighbor	-0.18		1383	-0.11	*	1383	0.246		1383	**0.01**	*	**1383**	**-0.28**	***	**808**	**0.407**	***	**1382**	0.108		1383	2.502		805	0.143		1383
PSM Stratification	-0.19		1467	-0.08		1467	-0.01		1467	**0.01**	*	**1467**	**-0.29**	***	**924**	**0.406**	**	**1467**	-0.11		1467	12.34		924	0.187		1467
PSW robust & cluster s.e.	-0.2		1442	-0.05		1442	-0.88	**	1439	**0.01**		**1442**	**-0.2**	***	**919**	**0.187**		**1438**	-0.07		1442	24.7		915	-0.08		1441
PSW Bootstrapped s.e.	-0.2		1442	-0.05		1442	-0.88	**	1427	**0.01**		**1442**	**-0.2**	***	**919**	**0.223**		**1438**	-0.07		1442	24.7		915	-0.08		1441

Methodology	Health																										
	Breathing difficulties in the Household (*Ordinal Categorical Variable*)			Yellowish skin in the Household (*Binary Categorical variable*)			Eyes disease symptoms in the Household (Binary Categorical *Variable*)			Gum disease symptoms in the Household (Binary Categorical *Variable*)			Beneficiary's Diarrhea symptoms (Ordinal Categorical Variable)			Beneficiary's Breathing difficulties (Ordinal Categorical Variable)			Beneficiary's Yellowish skin (Binary Categorical variable)			Beneficiary's Eyes disease symptoms (Binary Categorical Variable)			Beneficiary's Gum disease symptoms (Binary Categorical Variable)		
	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N
PSM Kernel	0.178		1467	0.002		1467	0.012		1467	0.024		1467	0.18		1467	0.139		1467	-0.02		1467	0.013		1467	0.002		1467
PSM Nearest Neighbor	0.129		1383	0		1383	-0.01		1383	0.018		1383	0.115		1383	0.026		1383	-0.01		1383	0.011		1383	0.002		1383
PSM Stratification	0.196		1467	0.004		1467	0.015		1467	0.026		1467	0.187		1467	0.141		1467	-0.02		1467	0.014		1467	0.002		1467
PSW robust & cluster s.e.	-0.08		1440	0.009		1440	0.008		1442	0.01		1442	-0.03		1442	-0.1		1439	9E-04		1442	0.02		1442	-0		1442
PSW Bootstrapped s.e.	-0.08		1440	0.009		1440	0.008		1442	0.01		1442	-0.03		1442	-0.1		1439	9E-04		1442	0.02		1442	-0		1442

Note: * is significant at the 90% confidence level, ** is significant at the 95% and *** is significant at the 99%. Estimations with bootstrapped standard errors are replicated 100 times. Robust standard errors are clustered at the locality level. The causal effect is a simple difference, since the response variables are only captured in a single point of time. The propensity score estimations include municipality fixed effects, while the structural equations consider locality fixed effects.

DEC Rural

Methodology	Food Support						Food Orientation																				
	Food insecurity by Household (*Continuous index*)			Food insecurity by Household (*Categorical index*)			Habit change perception on food selection by Household (Continuous Index)			Habit change perception on food selection by Household (Categorical Index)			Habit change perception on food preparation by Household (Continuous Index)			Habit change perception on food preparation by Household (Categorical Index)			Habit change perception on food consumption by Household (Continuous Index)			Habit change perception on food consumption by Household (Categorical Index)			Habit change perception by Household (Continuous Index)		
	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N
PSM Kernel	-0.15		957	-0.12		957	**5.636**	**	**957**	0.083		957	**3.702**	***	**957**	0.022		957	1.438		957	0.063		957	**4.158**	***	**957**
PSM Nearest Neighbor	0.163		745	0.061		745	**5.767**		**745**	0.155		745	**4.36**	*	**746**	-0.02		746	-0.1		742	0.027		742	**3.981**	*	**741**
PSM Stratification	-0.96		957	-0.05		957	**4.265**	*	**957**	0.085		957	**3.64**	**	**957**	0.018		957	1.063		957	0.065		957	**3.55**	***	**957**
PSW robust & cluster s.e.	-2.55	**	949	-0.17		949	**3.121**	**	**948**	0.001		948	**2.293**		**951**	0.087	***	951	-1.43		935	0.028		935	**1.198**		**932**
PSW Bootstrapped s.e.	-2.55	*	949	-0.17		949	**3.121**	*	**948**	0.025		948	**2.293**		**951**	0.087	***	951	-1.43		935	0.028		935	**1.198**		**932**

Methodology	Food Orientation																										
	Habit change perception by Household (Categorical Index)			Habit change perception on food selection by Beneficiary (*Continuous Index*)			Habit change perception on food selection by Beneficiary (*Categorical Index*)			Habit change perception on food consumption by Beneficiary (*Continuous Index*)			Habit change perception on food consumption by Beneficiary (*Categorical Index*)			Habit change perception by Beneficiary (*Continuous Index*)			Habit change perception by Beneficiary (*Categorical Index*)			Diet Diversity by Household (*Continuous Index*)			Diet Diversity by Household (*Continuous Index*)		
	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N
PSM Kernel	**0.184**	***	**957**	**10.17**	***	**957**	**0.225**	***	**957**	5.149		957	0.099		957	**8.8**	***	**957**	0.235	***	957	**-3.04**	**	**957**	-0.14	**	957
PSM Nearest Neighbor	**0.214**	*	**741**	**10.5**	***	**746**	**0.268**	***	**746**	2.406		738	0.05		738	**8.064**	***	**738**	0.128		738	**-2.9**	*	**746**	-0.11		746
PSM Stratification	**0.193**	***	**957**	**9.645**	***	**957**	**0.212**	***	**957**	3.975		957	0.083		957	**8.139**	***	**957**	0.182	*	957	**-3.4**	***	**957**	-0.14	**	957
PSW robust & cluster s.e.	**0.144**	**	**932**	**4.3**		**948**	**0.05**		**948**	-4.44	*	930	0.019		930	**0.919**		**927**	0.067		927	**-0.4**		**951**	-0.08		951
PSW Bootstrapped s.e.	**0.144**	*	**932**	**4.3**		**948**	**0.05**		**948**	-4.44		930	0.019		930	**0.919**		**927**	0.067		927	**-0.4**		**951**	-0.03		951

Methodology	Food Orientation												Education												Health		
	Diet Variety by Household (*Continuous Index*)			Diet Variety by Household (*Continuous Index*)			Diet Quality by Household (*Continuous Index*)			Diet Quality by Household (*Continuous Index*)			Student' Marks in Primary School			School Absence in last month (kinder & primary school)			School Absence in last schooling cycle (kinder & primary school)			Extra-curricular studies			Diarrhea symptoms in the Household (Ordinal Categorical Variable)		
	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N
PSM Kernel	**-0.93**	***	**957**	-0.1	**	957	-3.97	**	957	-0.02		957	-0.01		700	**0.453**	***	**957**	0.668	*	957	-27.9		700	-0.11		957
PSM Nearest Neighbor	**-0.9**	*	**746**	-0.11		746	3.792	**	746	-0.02		746	0.062		434	**0.548**	**	**746**	1.15	**	746	-26.7		432	0.069		746
PSM Stratification	**-0.91**	***	**957**	-0.11	**	957	-4.31	***	957	-0.02	*	957	0.033		700	**0.529**	***	**957**	0.637		957	-26.1		700	-0.03		957
PSW robust & cluster s.e.	**-0.39**		**951**	0.053		951	-0.79		951	-0.01		951	0.055		695	**0.456**	*	**946**	0.137		949	43.53		691	-0.08		951
PSW Bootstrapped s.e.	**-0.39**	*	**951**	0.053		951	-0.79	*	951	-0.01		951	0.055		695	**0.456**		**946**	0.137		946	43.53		691	-0.08		951

Methodology	Health																										
	Breathing difficulties in the Household (*Ordinal Categorical Variable*)			Yellowish skin in the Household (*Binary Categorical variable*)			Eyes disease symptoms in the Household (Binary Categorical *Variable*)			Gum disease symptoms in the Household (Binary Categorical *Variable*)			Beneficiary's Diarrhea symptoms (Ordinal Categorical Variable)			Beneficiary's Breathing difficulties (Ordinal Categorical Variable)			Beneficiary's Yellowish skin (Binary Categorical variable)			Beneficiary's Eyes disease symptoms (Binary Categorical Variable)			Beneficiary's Gum disease symptoms (Binary Categorical Variable)		
	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N
PSM Kernel	0.133		957	0.028		957	-0		957	0.009		957	-0.1		957	0.195		957	0.035		957	-0.01		957	-0.01		957
PSM Nearest Neighbor	0.332		746	0.039		746	0.023		746	0.01		746	0.026		746	0.444	**	746	0.046		746	-0.01		746	-0.02		746
PSM Stratification	0.172		957	0.023		957	0.028		957	-0.01		957	-0.04		957	0.233	*	957	0.028		957	0		957	-0.01		957
PSW robust & cluster s.e.	0.094		950	0.148	**	951	0.184	***	951	0.062		951	-0.07		950	0.265		950	0.131	**	951	0.106		951	0.05		951
PSW Bootstrapped s.e.	0.094		950	0.148		951	0.184	**	951	0.062		951	-0.07		950	0.265		950	0.131	**	951	0.106		951	0.05		951

Note: * is significant at the 90% confidence level, ** is significant at the 95% and *** is significant at the 99%. Estimations with bootstrapped standard errors are replicated 100 times. Robust standard errors are clustered at the locality level. The causal effect is a simple difference, since the response variables are only captured in a single point of time. The propensity score estimations include municipality fixed effects, while the structural equations consider locality fixed effects.

VIII.2. DEF Results by Sample

DEF General

	Food Support						Food Orientation																				
Methodology	Food insecurity by Household (*Continuous index*)			Food insecurity by Household (*Categorical index*)			Habit change perception on food selection by Household (Continuous Index)			Habit change perception on food selection by Household (Categorical Index)			Habit change perception on food preparation by Household (Continuous Index)			Habit change perception on food preparation by Household (Categorical Index)			Habit change perception on food consumption by Household (Continuous Index)			Habit change perception on food consumption by Household (Categorical Index)			Habit change perception by Household (Continuous Index)		
	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N
PSM Kernel	-0.16		1556	-0.11		1556	**4.203**	*	**1556**	0.06		1556	**2.994**	***	**1556**	**0.033**	*	**1556**	-0.79		1556	-0.01		1556	**2.618**	*	**1556**
PSM Nearest Neighbor	0.066		1295	-0.1		1295	**4.151**	**	**1292**	0.054		1292	**2.294**	*	**1295**	**0.034**		**1295**	-0.86		1284	-0.01		1284	**2.31**	*	**1281**
PSM Stratification	-0.28		1556	-0.11		1556	**4.134**		**1556**	0.062		1556	**2.933**	***	**1556**	**0.031**	*	**1556**	-0.78		1556	-0.01		1556	**2.584**	*	**1556**
PSW robust & cluster s.e.	-0.61		1539	-0.11		1539	**3.87**	*	**1529**	0.075		1529	**3.68**	***	**1540**	**0.043**	***	**1540**	-0.68		1507	-0.01		1507	**3.06**	*	**1494**
PSW Bootstrapped s.e.	-0.61		1539	-0.11		1539	**3.87**	*	**1529**	0.075		1529	**3.68**	***	**1540**	**0.043**	***	**1540**	-0.68		1507	-0.01		1507	**3.06**	*	**1494**

	Food Orientation																										
Methodology	Habit change perception by Household (Categorical Index)			Habit change perception on food selection by Beneficiary (*Continuous Index*)			Habit change perception on food selection by Beneficiary (*Categorical Index*)			Habit change perception on food consumption by Beneficiary (*Continuous Index*)			Habit change perception on food consumption by Beneficiary (*Categorical Index*)			Habit change perception by Beneficiary (*Continuous Index*)			Habit change perception by Beneficiary (*Categorical Index*)			Diet Diversity by Household (*Continuous Index*)			Diet Diversity by Household (*Continuous Index*)		
	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N
PSM Kernel	0.088		1556	**-8.84**	***	**1556**	**-0.22**	***	**1556**	**-7.14**	**	**1556**	**-0.19**	***	**1556**	**-8.97**	***	**1556**	**-0.28**	***	**1556**	-1.47	*	1556	-0.1	*	1556
PSM Nearest Neighbor	0.063		1281	**-9.35**	***	**1292**	**-0.23**	***	**1292**	**-6.19**	*	**1284**	**-0.16**	***	**1284**	**-8.92**	***	**1292**	**-0.28**	***	**1292**	-0.94		1295	-0.06		1295
PSM Stratification	0.084		1556	**-8.62**	***	**1556**	**-0.21**	***	**1556**	**-6.96**	**	**1556**	**-0.19**	***	**1556**	**-8.71**	***	**1556**	**-0.28**	***	**1556**	-1.53	*	1556	-0.1	*	1556
PSW robust & cluster s.e.	0.13	*	1494	**-3.25**		**1537**	**-0.1**		**1537**	**-9.95**	***	**1506**	**-0.25**	***	**1506**	**-6.26**	***	**1503**	**-0.24**	***	**1503**	-1.27		1541	-0.14		1541
PSW Bootstrapped s.e.	0.13	*	1494	**-3.25**		**1537**	**-0.1**		**1537**	**-9.95**	***	**1506**	**-0.25**	***	**1506**	**-6.26**	***	**1503**	**-0.24**	***	**1503**	-1.27		1541	-		1541

	Food Orientation												Education						Health								
Methodology	Diet Variety by Household (*Continuous Index*)			Diet Variety by Household (*Continuous Index*)			Diet Quality by Household (*Continuous Index*)			Diet Quality by Household (*Continuous Index*)			School Absence in last month (kinder & primary school)			School Absence in last schooling cycle (kinder & primary school)			Diarrhea symptoms in the Household (Ordinal Categorical Variable)			Breathing difficulties in the Household (*Ordinal Categorical Variable*)			Yellowish skin in the Household (*Binary Categorical variable*)		
	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N
PSM Kernel	-0.2		1556	-0.03		1556	-1.67	*	1556	-0.02		1556	**0.952**	**	**1556**	0.257		2426	-0.06		1556	0.012		1556	0.005		1556
PSM Nearest Neighbor	-0.18		1295	-0.03		1295	-1.12		1295	-0.04		1295	**0.933**	**	**1295**	0.352		2179	-0.03		1295	0.006		1295	-0		1295
PSM Stratification	-0.15		1556	-0.01		1556	-1.68	*	1556	0.00		1556	**0.898**	**	**1556**	0.294		2426	-0.08		1556	-0		1556	0.005		1556
PSW robust & cluster s.e.	0.57	***	1541	0.05		1541	-0.7		1541	-0.07	*	1541	**0.6**		**1540**	-0.07		2391	-0.3	**	1541	-0.38	***	1541	-0		1541
PSW Bootstrapped s.e.	0.57	***	1541	0.05		1541	-0.7		1541	-0.07	*	1541	**0.6**		**1540**	-0.07		2391	-0.3	**	1541	-0.38	***	1541	-0		1541

	Health																				
Methodology	Eyes disease symptoms in the Household (Binary Categorical *Variable*)			Gum disease symptoms in the Household (Binary Categorical *Variable*)			Beneficiary's Diarrhea symptoms (Ordinal Categorical Variable)			Beneficiary's Breathing difficulties (Ordinal Categorical Variable)			Beneficiary's Yellowish skin (Binary Categorical variable)			Beneficiary's Eyes disease symptoms (Binary Categorical Variable)			Beneficiary's Gum disease symptoms (Binary Categorical Variable)		
	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N
PSM Kernel	0		1556	0.005		1556	-0.05		1556	0		1556	0.008		1556	-0.01		1556	0.002		1556
PSM Nearest Neighbor	-0		1295	0.019		1295	-0.02		1294	-0.09		1295	-0		1295	-0.02		1295	0.009		1295
PSM Stratification	0.002		1556	0.008		1556	-0.05		1556	-0.03		1556	0.009		1556	-0.01		1556	0.002		1556
PSW robust & cluster s.e.	-0.01		1541	-0.05	**	1541	-0.27	*	1539	-0.21	**	1539	-0.02		1541	-0.01		1541	-0.03	**	1540
PSW Bootstrapped s.e.	-0.01		1541	-0.05	**	1541	-0.27	*	1539	-0.21	***	1539	-0.02		1541	-0.01		1541	-0.03	**	1540

Note: * is significant at the 90% confidence level, ** is significant at the 95% and *** is significant at the 99%. Estimations with bootstrapped standard errors are replicated 100 times. Robust standard errors are clustered at the locality level. The causal effect is a simple difference, since the response variables are only captured in a single point of time. The propensity score estimations include municipality fixed effects, while the structural equations consider locality fixed effects.

DEF Boys

	Food Support						Food Orientation																				
Methodology	Food insecurity by Household (*Continuous index*)			Food insecurity by Household (*Categorical index*)			Habit change perception on food <u>selection</u> by Household (Continuous Index)			Habit change perception on food <u>selection</u> by Household (Categorical Index)			Habit change perception on food <u>preparation</u> by Household (Continuous Index)			Habit change perception on food <u>preparation</u> by Household (Categorical Index)			Habit change perception on food <u>consumption</u> by Household (Continuous Index)			Habit change perception on food <u>consumption</u> by Household (Categorical Index)			Habit change perception by Household (Continuous Index)		
	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N
PSM Kernel	-0.56		779	-0.13		779	3.851		779	0.033		779	1.35		779	0.026		779	-1.96		779	-0.04		779	1.158		779
PSM Nearest Neighbor	-0.45		643	-0.1		643	4.077		642	0.071		642	1.023		643	0.026		643	-1.69		638	-0.04		638	1.255		637
PSM Stratification	-0.71		779	-0.14		779	3.566		779	0.025		779	1.25		779	0.022		779	-2.1		779	-0.05		779	0.981		779
PSW robust & cluster s.e.	-0.29		776	-0.06		776	1.88		769	0.004		769	0.28		776	0.25		776	-2.74	*	761	-0.07	*	761	-0.2		752
PSW Bootstrapped s.e.	-0.29		776	-0.06		776	1.88		769	0.004		769	0.28		776	0.25		776	-2.74	*	761	-0.07	*	761	-0.2		752

	Food Orientation																										
Methodology	Habit change perception by Household (Categorical Index)			Habit change perception on food <u>selection</u> by Beneficiary (*Continuous Index*)			Habit change perception on food <u>selection</u> by Beneficiary (*Categorical Index*)			Habit change perception on food <u>consumption</u> by Beneficiary (*Continuous Index*)			Habit change perception on food <u>consumption</u> by Beneficiary (*Categorical Index*)			Habit change perception by Beneficiary (*Continuous Index*)			Habit change perception by Beneficiary (*Categorical Index*)			Diet <u>Diversity</u> by Household (*Continuous Index*)			Diet <u>Diversity</u> by Household (*Continuous Index*)		
	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N
PSM Kernel	0.053		779	**-8.99**	***	**779**	**-0.19**	***	**779**	**-9.09**	**	**779**	**-0.23**	***	**779**	**-9.82**	***	**779**	**-0.34**	***	**779**	-1.04		779	-0.08		779
PSM Nearest Neighbor	0.099	*	637	**-9.99**	***	**642**	**-0.2**	***	**642**	**-8.81**	**	**638**	**-0.2**	***	**638**	**-10.4**	***	**638**	**-0.36**	***	**638**	-0.79		643	-0.08		643
PSM Stratification	0.043		779	**-9.12**	*	**779**	**-0.19**	***	**779**	**-9.11**	**	**779**	**-0.23**	***	**779**	**-9.85**	***	**779**	**-0.35**	***	**779**	-1.13		779	-0.08		779
PSW robust & cluster s.e.	0.01		752	**-5.41**	*	**776**	**-0.11**		**776**	**-11.6**	***	**762**	**-0.27**	***	**762**	**-8.35**	***	**762**	**-0.35**	***	**762**	-0.46		777	-0.1		777
PSW Bootstrapped s.e.	0.01		752	**-5.41**	*	**776**	**-0.11**		**776**	**-11.6**	***	**762**	**-0.27**	***	**762**	**-8.35**	***	**762**	**-0.35**	***	**762**	-0.46		777	-0.1		777

	Food Orientation												Education						Health								
Methodology	Diet <u>Variety</u> by Household (*Continuous Index*)			Diet <u>Variety</u> by Household (*Continuous Index*)			Diet <u>Quality</u> by Household (*Continuous Index*)			Diet <u>Quality</u> by Household (*Continuous Index*)			School Absence in last month (kinder & primary school)			School Absence in last schooling cycle (kinder & primary school)			<u>Diarrhea</u> symptoms in the Household (Ordinal Categorical Variable)			<u>Breathing</u> difficulties in the Household *(Ordinal Categorical Variable)*			<u>Yellowish skin</u> in the Household (*Binary Categorical variable*)		
	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N
PSM Kernel	-0.55		779	-0.07		779	-1.58		779	-0		779	0.101		779	0.878		779	0.097		779	0.128		779	0.015		779
PSM Nearest Neighbor	-0.57		643	-0.07		643	-1.36		643	0.001		643	0.186		643	0.753		643	0.122		643	0.096		643	0.01		643
PSM Stratification	-0.51		779	-0.06		779	-1.64		779	-0		779	0.098		779	0.832		779	0.098		779	0.107		779	0.016		779
PSW robust & cluster s.e.	-0.02		777	-0		777	-0.48		777	-0.01	*	777	0.25		776	0.73		777	-0.23		777	-0.22		776	-0		777
PSW Bootstrapped s.e.	-0.02		777	-0		777	-0.48		777	-0.01		777	0.25		776	+		777	-0.23		777	-0.22		776	-0		777

	Health																				
Methodology	<u>Eyes</u> disease symptoms in the Household (Binary Categorical *Variable*)			<u>Gum</u> disease symptoms in the Household (Binary Categorical *Variable*)			Beneficiary's <u>Diarrhea</u> symptoms (Ordinal Categorical Variable)			Beneficiary's <u>Breathing</u> difficulties (Ordinal Categorical Variable)			Beneficiary's <u>Yellowish skin</u> (Binary Categorical variable)			Beneficiary's <u>Eyes</u> disease symptoms (Binary Categorical Variable)			Beneficiary's <u>Gum</u> disease symptoms (Binary Categorical Variable)		
	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N
PSM Kernel	-0		779	-0.01		779	0.07		779	0.072		779	0.012		779	-0.01		779	0.009		779
PSM Nearest Neighbor	-0		643	0.016		643	0.121		643	-0.02		643	-0		643	-0.01		643	0.016		643
PSM Stratification	-0.01		779	-0		779	0.074		779	0.054		779	0.012		779	-0.09		779	0.009		779
PSW robust & cluster s.e.	0.026		777	-0.11	**	777	-0.12		776	-0.09		777	0.001		777	-0.02		777	-0.01		776
PSW Bootstrapped s.e.	0.026		777	-0.11	**	777	-0.12		776	-0.09		777	0.001		777	-0.02		777	-0.01		776

Note: * is significant at the 90% confidence level, ** is significant at the 95% and *** is significant at the 99%. Estimations with bootstrapped standard errors are replicated 100 times. Robust standard errors are clustered at the locality level. The causal effect is a simple difference, since the response variables are only captured in a single point of time. The propensity score estimations include municipality fixed effects, while the structural equations consider locality fixed effects.

DEF Girls

	Food Support						Food Orientation																				
Metodología	Food insecurity by Household (*Continuous index*)			Food insecurity by Household (*Categorical index*)			Habit change perception on food <u>selection</u> by Household (Continuous Index)			Habit change perception on food <u>selection</u> by Household (Categorical Index)			Habit change perception on food <u>preparation</u> by Household (Continuous Index)			Habit change perception on food <u>preparation</u> by Household (Categorical Index)			Habit change perception on food <u>consumption</u> by Household (Continuous Index)			Habit change perception on food <u>consumption</u> by Household (Categorical Index)			Habit change perception by Household (Continuous Index)		
	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N
PSM Kernel	0.309		773	-0.09		773	**4.331**	**	**773**	0.082		773	**4.58**	***	**773**	**0.034**	*	**773**	0.608		773	0.036		773	**4.071**	***	**773**
PSM Nearest Neighbor	1.45		589	-0.05		589	**2.2**		**586**	-0		586	**4.588**	**	**589**	**0.01**		**589**	1.55		586	0.046		586	**3.837**	***	**583**
PSM Stratification	0.266		773	-0.09		773	**4.339**	**	**773**	0.113		773	**4.584**	***	**773**	**0.032**	*	**773**	0.99		773	0.023		773	**4.25**	***	**773**
PSW robust & cluster s.e.	0.35		763	-0.08		763	**5.58**	**	**760**	0.145	*	760	**6.29**	***	**764**	**0.045**	**	**764**	1.23		746	0.05		746	**5.92**	**	**742**
PSW Bootstrapped s.e.	0.35		763	-0.08		763	**5.58**	*	**760**	0.145		760	**6.29**	***	**764**	**0.045**	*	**764**	1.23		746	0.05		746	**5.92**	**	**742**

	Food Orientation																										
Methodology	Habit change perception by Household (Categorical Index)			Habit change perception on food <u>selection</u> by Beneficiary (*Continuous Index*)			Habit change perception on food <u>selection</u> by Beneficiary (*Categorical Index*)			Habit change perception on food <u>consumption</u> by Beneficiary (*Continuous Index*)			Habit change perception on food <u>consumption</u> by Beneficiary (*Categorical Index*)			Habit change perception by Beneficiary (*Continuous Index*)			Habit change perception by Beneficiary (*Categorical Index*)			Diet <u>Diversity</u> by Household (*Continuous Index*)			Diet <u>Diversity</u> by Household (*Continuous Index*)		
	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N
PSM Kernel	**0.12**	***	**773**	**-9.1**	***	**773**	**-0.27**	***	**773**	-5.09		773	-0.16	*	773	**-8.35**	***	**773**	**-0.22**	***	**773**	**-1.89**	***	**773**	**-0.12**	***	**773**
PSM Nearest Neighbor	**0.063**		**583**	**-12.3**	***	**588**	**-0.34**	***	**588**	-1.44		585	-0.1		585	**-8.69**	***	**584**	**-0.18**	**	**584**	**-0.53**		**589**	**-0.06**		**589**
PSM Stratification	**0.117**	***	**773**	**-8.78**	***	**773**	**-0.27**	***	**773**	-6.29		773	-0.18	*	773	**-8.6**	***	**773**	**-0.23**	***	**773**	**-0.17**	***	**773**	**-0.11**	***	**773**
PSW robust & cluster s.e.	**0.198**	***	**742**	**-2.25**		**761**	**-0.12**	**	**761**	-8.27		744	-0.23		744	**-4.96**		**741**	**-0.13**		**741**	**-1.65**	**	**764**	**-0.16**	***	**764**
PSW Bootstrapped s.e.	**0.198**	***	**742**	**-2.25**		**761**	**-0.12**	*	**761**	-8.27		744	-0.23		744	**-4.96**		**741**	**-0.13**		**741**	**-1.65**	**	**764**	**-0.16**	**	**764**

	Food Orientation												Education						Health								
Methodology	Diet <u>Variety</u> by Household (*Continuous Index*)			Diet <u>Variety</u> by Household (*Continuous Index*)			Diet <u>Quality</u> by Household (*Continuous Index*)			Diet <u>Quality</u> by Household (*Continuous Index*)			School Absence in last month (kinder & primary school)			School Absence in last schooling cycle (kinder & primary school)			<u>Diarrhea</u> symptoms in the Household (Ordinal Categorical Variable)			<u>Breathing</u> difficulties in the Household *(Ordinal Categorical Variable)*			<u>Yellowish skin</u> in the Household (*Binary Categorical variable*)		
	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N
PSM Kernel	0.146		773	0.017		773	-1.74	**	773	-0		773	0.062		773	1.002	*	773	**-0.21**	**	**773**	-0.1		773	-0		773
PSM Nearest Neighbor	0.213		589	0.052		589	-0.32		589	0		589	0.155		589	0.929		589	**-0.25**	*	**589**	-0.06		589	0.002		589
PSM Stratification	0.282		773	0.062		773	-1.39	*	773	-0		773	0.021		773	0.915	*	773	**-0.35**	***	**773**	-0.15		773	-0.01		773
PSW robust & cluster s.e.	1.23	**	764	0.121	**	764	-0.42		764	-0		764	-0.2		763	0.57		763	**-0.43**	***	**764**	-0.54	***	764	-0.01		764
PSW Bootstrapped s.e.	1.23	*	764	0.121	**	764	-0.42		764	-0		764	-0.2		763	0.57		763	**-0.43**	***	**764**	-0.54	***	764	-0.01		764

	Health																				
Methodology	<u>Eyes</u> disease symptoms in the Household (Binary Categorical *Variable*)			<u>Gum</u> disease symptoms in the Household (Binary Categorical *Variable*)			Beneficiary's <u>Diarrhea</u> symptoms (Ordinal Categorical Variable)			Beneficiary's <u>Breathing</u> difficulties (Ordinal Categorical Variable)			Beneficiary's <u>Yellowish skin</u> (Binary Categorical variable)			Beneficiary's <u>Eyes</u> disease symptoms (Binary Categorical Variable)			Beneficiary's <u>Gum</u> disease symptoms (Binary Categorical Variable)		
	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N
PSM Kernel	0.001		773	0.019		773	**-0.18**		**773**	-0.07		773	0.006		773	-0		773	-0.01		773
PSM Nearest Neighbor	0.019		589	0.032		589	**-0.13**		**588**	-0.16		589	0.019		589	0.008		589	0.017		589
PSM Stratification	0.007		773	0.019		773	**-0.25**	**	**773**	-0.19		773	0.006		773	-0		773	-0.01		773
PSW robust & cluster s.e.	-0.04	*	764	-0.01		764	**-0.46**	***	**763**	-0.36	***	762	-0.05	*	764	-0.02		764	-0.05		764
PSW Bootstrapped s.e.	-0.04		764	-0.01		764	**-0.46**	***	**763**	-0.36	***	762	-0.05		764	-0.02		764	-0.05		764

Note: * is significant at the 90% confidence level, ** is significant at the 95% and *** is significant at the 99%. Estimations with bootstrapped standard errors are replicated 100 times. Robust standard errors are clustered at the locality level. The causal effect is a simple difference, since the response variables are only captured in a single point of time. The propensity score estimations include municipality fixed effects, while the structural equations consider locality fixed effects.

DEF Urban

	Food Support						Food Orientation																				
Methodology	Food insecurity by Household (*Continuous index*)			Food insecurity by Household (*Categorical index*)			Habit change perception on food selection by Household (Continuous Index)			Habit change perception on food selection by Household (Categorical Index)			Habit change perception on food preparation by Household (Continuous Index)			Habit change perception on food preparation by Household (Categorical Index)			Habit change perception on food consumption by Household (Continuous Index)			Habit change perception on food consumption by Household (Categorical Index)			Habit change perception by Household (Continuous Index)		
	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N
PSM Kernel	0.068		1024	-0.1		1024	**5.326**	**	**1024**	0.105		1024	**3.84**	***	**1024**	0.035		1024	-1.15		1024	-0.02		1024	**3.24**	**	**1024**
PSM Nearest Neighbor	0.136		941	-0.09		941	**4.593**	*	**939**	0.089		939	**2.87**	**	**941**	0.031		941	-1.61		936	-0.04		936	**2.43**	*	**934**
PSM Stratification	-0.08		1024	-0.1		1024	**5.321**	*	**1024**	0.11		1024	**3.78**	***	**1024**	0.037		1024	-1.22		1024	-0.03		1024	**3.19**	**	**1024**
PSW robust & cluster s.e.	-0.69		1011	-0.12		1011	**3.75**		**1003**	0.08		1003	**3.84**	***	**1011**	0.038	***	1011	-0.55		994	-0.01		994	**3.2**	**	**904**
PSW Bootstrapped s.e.	-0.69		1011	-0.12		1011	**3.75**		**1003**	0.08		1003	**3.84**	***	**1011**	0.038	**	1011	-0.55		994	-0.01		994	**3.2**	*	**904**

	Food Orientation																										
Methodology	Habit change perception by Household (Categorical Index)			Habit change perception on food selection by Beneficiary (*Continuous Index*)			Habit change perception on food selection by Beneficiary (*Categorical Index*)			Habit change perception on food consumption by Beneficiary (*Continuous Index*)			Habit change perception on food consumption by Beneficiary (*Categorical Index*)			Habit change perception by Beneficiary (*Continuous Index*)			Habit change perception by Beneficiary (*Categorical Index*)			Diet Diversity by Household (*Continuous Index*)			Diet Diversity by Household (*Continuous Index*)		
	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N
PSM Kernel	**0.105**	*	**1024**	**-7.7**	***	**1024**	**-0.19**	***	**1024**	**-8.2**	**	**1024**	**-0.2**	**	**1024**	**-8.71**	***	**1024**	**-0.28**	***	**1024**	-1.78	*	1024	**-0.12**	**	**1024**
PSM Nearest Neighbor	**0.066**		**934**	**-9.17**	***	**938**	**-0.22**	***	**938**	**-7.02**	*	**935**	**-0.17**	**	**935**	**-9.22**	***	**933**	**-0.29**	***	**933**	-1.29		941	**-0.08**	*	**941**
PSM Stratification	**0.101**	*	**1024**	**-7.42**	***	**1024**	**-0.18**	***	**1024**	**-7.93**	**	**1024**	**-0.19**	***	**1024**	**-8.38**	***	**1024**	**-0.28**	***	**1024**	-1.84	*	1024	**-0.12**	**	**1024**
PSW robust & cluster s.e.	**0.13**	*	**904**	**-3.53**		**1009**	**-1.04**		**1011**	**-10.3**	***	**992**	**-0.26**	***	**992**	**-6.57**	***	**990**	**-0.25**	***	**990**	-1.16		1012	**-0.13**		**1012**
PSW Bootstrapped s.e.	**0.13**	*	**904**	**-3.53**		**1011**	**-1.04**		**1011**	**-10.3**	***	**992**	**-0.26**	***	**992**	**-6.57**	**	**990**	**-0.25**	***	**990**	-1.16		1012	**-0.13**		**1012**

	Food Orientation												Education						Health								
Methodology	Diet Variety by Household (*Continuous Index*)			Diet Variety by Household (*Continuous Index*)			Diet Quality by Household (*Continuous Index*)			Diet Quality by Household (*Continuous Index*)			School Absence in last month (kinder & primary school)			School Absence in last schooling cycle (kinder & primary school)			Diarrhea symptoms in the Household (Ordinal Categorical Variable)			Breathing difficulties in the Household (*Ordinal Categorical Variable*)			Yellowish skin in the Household (*Binary Categorical variable*)		
	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N
PSM Kernel	-0.16		1024	-0.02		1024	-1.94	*	1024	**-0.01**	**	**1024**	0.173		1024	0.83	*	1024	-0.09		1024	0.006		1024	0.007		1024
PSM Nearest Neighbor	-0.04		941	0.004		941	-1.33		941	**-0.01**	**	**941**	0.138		941	0.713		941	-0.08		941	-0.02		941	-0.01		941
PSM Stratification	-0.09		1024	-0.01		1024	-1.93	*	1024	**-0.01**	**	**1024**	0.164		1024	0.776		1024	-0.11		1024	-0.01		1024	0.006		1024
PSW robust & cluster s.e.	0.59	***	1012	0.057		1012	-0.57		1012	**-0.01**	*	**1012**	0.026		1011	0.6		1012	-0.28	**	1012	-0.37	***	1012	-0		1012
PSW Bootstrapped s.e.	0.59	***	1012	0.057		1012	-0.57		1012	**-0.01**		**1012**	0.026		1011	0.6		1012	-0.28	*	1012	-0.37	***	1012	-0		1012

	Health																				
Methodology	Eyes disease symptoms in the Household (Binary Categorical *Variable*)			Gum disease symptoms in the Household (Binary Categorical *Variable*)			Beneficiary's Diarrhea symptoms (Ordinal Categorical Variable)			Beneficiary's Breathing difficulties (Ordinal Categorical Variable)			Beneficiary's Yellowish skin (Binary Categorical variable)			Beneficiary's Eyes disease symptoms (Binary Categorical Variable)			Beneficiary's Gum disease symptoms (Binary Categorical Variable)		
	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N
PSM Kernel	0.013		1024	-0		1024	-0.09		1024	-0.02		1024	0.006		1024	-0.01		1024	-0.01		1024
PSM Nearest Neighbor	0.005		941	0.016		941	-0.06		940	-0.13		941	-0.01		941	-0.02		941	-0		941
PSM Stratification	0.015		1024	0		1024	-0.09		1024	-0.04		1024	0.005		1024	-0.01		1024	-0.01		1024
PSW robust & cluster s.e.	-0.01		1012	-0.05	**	1012	-0.29	*	1010	-0.21	**	1010	-0.02		1012	-0.01		1012	-0.03	**	1011
PSW Bootstrapped s.e.	-0.01		1012	-0.05	**	1012	-0.29	*	1010	-0.21	**	1010	-0.02		1012	-0.01		1012	-0.03	**	1011

Note: * is significant at the 90% confidence level, ** is significant at the 95% and *** is significant at the 99%. Estimations with bootstrapped standard errors are replicated 100 times. Robust standard errors are clustered at the locality level. The causal effect is a simple difference, since the response variables are only captured in a single point of time. The propensity score estimations include municipality fixed effects, while the structural equations consider locality fixed effects.

DEF Rural

	Food Support						Food Orientation																				
Methodology	Food insecurity by Household (*Continuous index*)			Food insecurity by Household (*Categorical index*)			Habit change perception on food selection by Household (Continuous Index)			Habit change perception on food selection by Household (Categorical Index)			Habit change perception on food preparation by Household (Continuous Index)			Habit change perception on food preparation by Household (Categorical Index)			Habit change perception on food consumption by Household (Continuous Index)			Habit change perception on food consumption by Household (Categorical Index)			Habit change perception by Household (Continuous Index)		
	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N
PSM Kernel	-0.72		532	-0.12	*	532	-0.47		532	-0.13		532	-0.15		532	0.022		532	0.35		532	0.034		532	0.089		532
PSM Nearest Neighbor	0.504		358	-0.04		358	0.074		357	-0.1		357	-0.28		358	0.005		358	2.49		352	0.067		352	1.105		351
PSM Stratification	-0.96		532	-0.14	*	532	-0.17		532	-0.12		532	-0.1		532	0.016		532	0.57		532	0.045		532	0.31		532
PSW robust & cluster s.e.	1.83		528	0.34		528	4.54	**	526	0.075		526	-8.73	***	529	0.17		529	3.2	*	513	0.26	***	513	-2.05		510
PSW Bootstrapped s.e.	1.83		528	0.34		528	4.54		526	0.075		526	-8.73	*	529	0.17		529	3.2		513	0.26		513	-2.05		510

	Food Orientation																										
Methodology	Habit change perception by Household (Categorical Index)			Habit change perception on food selection by Beneficiary (*Continuous Index*)			Habit change perception on food selection by Beneficiary (*Categorical Index*)			Habit change perception on food consumption by Beneficiary (*Continuous Index*)			Habit change perception on food consumption by Beneficiary (*Categorical Index*)			Habit change perception by Beneficiary (*Continuous Index*)			Habit change perception by Beneficiary (*Categorical Index*)			Diet Diversity by Household (*Continuous Index*)			Diet Diversity by Household (*Continuous Index*)		
	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N
PSM Kernel	0.02		532	**-13.5**	***	**532**	**-0.37**	***	**532**	-3.03		532	-0.16		532	-10.1	*	532	**-0.27**	**	**532**	0.043		532	0.004		532
PSM Nearest Neighbor	0.012		351	**-12.5**	***	**358**	**-0.35**	***	**358**	1.433		353	-0.07		353	-7.52		353	**-0.22**	*	**353**	0.876		358	0.076		358
PSM Stratification	0.017		532	**-13.1**	***	**532**	**-0.36**	***	**532**	-3.32		532	-0.16		532	-9.99	*	532	**-0.28**	**	**532**	-0.15		532	-0.01		532
PSW robust & cluster s.e.	-0.14		510	**0.04**		**528**	**0.107**		**528**	-10.2	*	514	-0.05		514	-4.11		513	**-0.04**		**513**	2.03	**	529	0.104		529
PSW Bootstrapped s.e.	-0.14		510	**0.04**		**528**	**0.107**		**528**	-10.2		514	-0.05		514	-4.11		513	**-0.04**		**513**	2.03		529	0.104		529

	Food Orientation												Education						Health								
Methodology	Diet Variety by Household (*Continuous Index*)			Diet Variety by Household (*Continuous Index*)			Diet Quality by Household (*Continuous Index*)			Diet Quality by Household (*Continuous Index*)			School Absence in last month (kinder & primary school)			School Absence in last schooling cycle (kinder & primary school)			Diarrhea symptoms in the Household (Ordinal Categorical Variable)			Breathing difficulties in the Household (*Ordinal Categorical Variable*)			Yellowish skin in the Household (*Binary Categorical variable*)		
	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N
PSM Kernel	-0.41		532	-0.05		532	-0.37		532	0.017		532	-0.3	*	532	1.46		532	0.095		532	0.03		532	-0		532
PSM Nearest Neighbor	-0.79		358	-0.13		358	0.091		358	0.02	*	358	-0.44		358	1.61	*	358	0.109		358	0.036		358	0.014		358
PSM Stratification	-0.38		532	-0.04		532	-0.53		532	0.017		532	-0.25		532	1.47		532	0.07		532	0.053		532	0.001		532
PSW robust & cluster s.e.	-0.1		529	-0.12		529	1.93	*	529	-0		529	-0.53		528	-2.47		528	-0.05		529	-0.1		529	0.016		529
PSW Bootstrapped s.e.	-0.1		529	-0.12		529	1.93		529	-0		529	-0.53		528	-2.47		528	-0.05		529	-0.1		529	0.016		529

	Health																				
Methodology	Eyes disease symptoms in the Household (Binary Categorical *Variable*)			Gum disease symptoms in the Household (Binary Categorical *Variable*)			Beneficiary's Diarrhea symptoms (Ordinal Categorical Variable)			Beneficiary's Breathing difficulties (Ordinal Categorical Variable)			Beneficiary's Yellowish skin (Binary Categorical variable)			Beneficiary's Eyes disease symptoms (Binary Categorical Variable)			Beneficiary's Gum disease symptoms (Binary Categorical Variable)		
	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N
PSM Kernel	-0.06	*	532	0.034		532	0.117		532	0.062		532	0.021		532	-0.01		532	0.052		532
PSM Nearest Neighbor	-0.04		358	0.057	*	358	0.064		358	0.046		358	0.021		358	0		358	0.069	**	358
PSM Stratification	-0.05		532	0.036		532	0.098		532	0.062		532	0.019		532	-0		532	0.058	*	532
PSW robust & cluster s.e.	-0.03		529	-0		529	0.14		529	-0.09		529	-0.07		529	-0.05		529	-0.04		529
PSW Bootstrapped s.e.	-0.03		529	-0		529	0.14		529	-0.09		529	-0.07		529	-0.05		529	-0.04		529

Note: * is significant at the 90% confidence level, ** is significant at the 95% and *** is significant at the 99%. Estimations with bootstrapped standard errors are replicated 100 times. Robust standard errors are clustered at the locality level. The causal effect is a simple difference, since the response variables are only captured in a single point of time. The propensity score estimations include municipality fixed effects, while the structural equations consider locality fixed effects.

VIII.3. INC Results by Sample

INC General

Methodology	Food Support												Food Orientation														
	WAZ			HAZ			WHZ			BMI			Habit change perception on food selection by Household (Continuous Index)			Habit change perception on food selection by Household (Categorical Index)			Habit change perception on food preparation by Household (Continuous Index)			Habit change perception on food preparation by Household (Categorical Index)			Habit change perception on food consumption by Household (Continuous Index)		
	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N
PSM Kernel	**0.26**	***	**1955**	**0.35**	***	**1955**	0.108		1955	0.046		1955	0.97		1955	0.019		1955	0.575		1955	0.00		1955	0.093		1955
PSM Nearest Neighbor	**0.20**	***	**1854**	**0.16**		**1854**	0.18		1854	0.158		1854	1.07		1849	0.037		1849	0.331		1854	-0.01		1854	-0.1		1844
PSM Stratification	**0.25**	***	**1955**	**0.35**	**	**1955**	0.09		1955	0.026		1955	0.66		1955	0.017		1955	0.256		1955	-0.01		1955	-0.26		1955
PSW robust & cluster s.e.	**0.29**	***	**1942**	**0.44**	***	**1942**	0.06		1939	0.0		1942	0.70		1934	0.02		1934	0.38		1942	0.00		1942	-0.12		1904
PSW Bootstrapped s.e.	**0.29**	***	**1942**	**0.44**	***	**1942**	0.06		1939	0.0		1942	0.70		1934	0.02		1934	0.38		1942	0.00		1942	-0.12		1904

Methodology	Food Orientation																										
	Habit change perception on food consumption by Household (Categorical Index)			Habit change perception by Household (Continuous Index)			Habit change perception by Household (Categorical Index)			Habit change perception on food selection by Beneficiary (*Continuous Index*)			Habit change perception on food selection by Beneficiary (*Categorical Index*)			Habit change perception on food consumption by Beneficiary (*Continuous Index*)			Habit change perception on food consumption by Beneficiary (*Categorical Index*)			Habit change perception by Beneficiary (*Continuous Index*)			Habit change perception by Beneficiary (*Categorical Index*)		
	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N
PSM Kernel	0.039		1955	0.519		1955	0.049	*	1955	**1.78**	*	**1955**	**0.08**	**	**1955**	-0.179		1955	-0.02		1955	1.185		1955	0.01		1955
PSM Nearest Neighbor	0.023		1844	0.35		1839	0.006		1839	**2.12**	*	**1854**	**0.09**	**	**1854**	-1.24		1846	-0.03		1846	0.946		1846	0.00		1846
PSM Stratification	0.027		1955	0.173		1955	0.032		1955	**1.74**		**1955**	**0.09**	**	**1955**	-0.797		1955	-0.04		1955	0.894		1955	0.00		1955
PSW robust & cluster s.e.	0.03		1904	0.18		1896	0.032	*	1896	**1.83**	**	**1939**	**0.09**	***	**1939**	-1.85		1907	-0.05	*	1907	0.39		1904	-0.01		1904
PSW Bootstrapped s.e.	0.03		1904	0.18		1896	0.032		1896	**1.83**	*	**1939**	**0.09**	***	**1939**	-1.85		1907	-0.05		1907	0.39		1904	-0.01		1904

Methodology	Food Orientation																		Health								
	Diet Diversity by Household (*Continuous Index*)			Diet Diversity by Household (*Continuous Index*)			Diet Variety by Household (*Continuous Index*)			Diet Variety by Household (*Continuous Index*)			Diet Quality by Household (*Continuous Index*)			Diet Quality by Household (*Continuous Index*)			Diarrhea symptoms in the Household (Ordinal Categorical Variable)			Breathing difficulties in the Household (*Ordinal Categorical Variable*)			Yellowish skin in the Household (*Binary Categorical variable*)		
	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N
PSM Kernel	-0.56		1955	-0.02		1955	-0.21		1955	0.00		1955	-0.77		1955	**-0.01**	*	**1955**	-0.04		1955	0.0		1955	0.0		1955
PSM Nearest Neighbor	-0.43		1854	0.00		1854	-0.23		1854	-0.01		1854	-0.66		1854	**-0.01**		**1854**	-0.04		1854	0.0		1854	0.0		1854
PSM Stratification	-0.39		1955	-0.01		1955	-0.19		1955	0.00		1955	-0.58		1955	**0.00**	*	**1955**	-0.03		1955	0.0		1955	0.0		1955
PSW robust & cluster s.e.	-0.23		1942	0.00		1942	-0.22		1942	-0.01		1942	-0.45		1942	**-0.01**	*	**1942**	-0.01		1942	0.0		1942	0.0		1942
PSW Bootstrapped s.e.	-0.23		1942	0.00		1942	-0.22		1942	-0.01		1942	-0.45		1942	**-0.01**		**1942**	-0.01		1942	0.0		1942	0.0		1942

Methodology	Health																				
	Eyes disease symptoms in the Household (Binary Categorical *Variable*)			Gum disease symptoms in the Household (Binary Categorical *Variable*)			Beneficiary's Diarrhea symptoms (Ordinal Categorical Variable)			Beneficiary's Breathing difficulties (Ordinal Categorical Variable)			Beneficiary's Yellowish skin (Binary Categorical variable)			Beneficiary's Eyes disease symptoms (Binary Categorical Variable)			Beneficiary's Gum disease symptoms (Binary Categorical Variable)		
	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N
PSM Kernel	-0.01		1955	-0.01		1955	0.023		1955	0.016		1955	-0.01		1955	0.00		1955	**-0.01**		**1955**
PSM Nearest Neighbor	-0.01		1854	-0.022		1854	0.012		1854	0.053		1854	-0.01		1854	0.00		1854	**-0.02**		**1854**
PSM Stratification	0.00		1955	-0.007		1955	0.022		1955	0.03		1955	-0.01		1955	0.00		1955	**-0.01**	*	**1955**
PSW robust & cluster s.e.	0.00		1942	-0.006		1942	0.04		1942	0.015		1942	0.00		1941	0.00		1942	**-0.02**	*	**1942**
PSW Bootstrapped s.e.	0.00		1942	-0.006		1942	0.04		1942	0.015		1942	0.00		1941	0.00		1942	**-0.02**	*	**1942**

Note: * is significant at the 90% confidence level, ** is significant at the 95% and *** is significant at the 99%. Estimations with bootstrapped standard errors are replicated 100 times. Robust standard errors are clustered at the locality level. The causal effect is a simple difference, since the response variables are only captured in a single point of time. The propensity score estimations include municipality fixed effects, while the structural equations consider locality fixed effects.

INC Boys

	Food Support												Food Orientation														
Methodology	WAZ			HAZ			WHZ			BMI			Habit change perception on food <u>selection</u> by Household (Continuous Index)			Habit change perception on food <u>selection</u> by Household (Categorical Index)			Habit change perception on food <u>preparation</u> by Household (Continuous Index)			Habit change perception on food <u>preparation</u> by Household (Categorical Index)			Habit change perception on food <u>consumption</u> by Household (Continuous Index)		
	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N
PSM Kernel	0.214	**	1010	0.327	**	1010	0.029		1010	-0.02		1010	0.992		1010	0.042		1010	-0.13		1010	-0.01		1010	-0.49		1010
PSM Nearest Neighbor	0.183		936	0.249	*	936	0.042		936	0.012		936	-0.54		933	0.007		933	-0.46		936	-0.01		936	-1.4		933
PSM Stratification	0.201	*	1010	0.311	**	1010	0.022		1010	-0.03		1010	0.505		1010	0.042		1010	-0.4		1010	-0.02		1010	-0.91		1010
PSW robust & cluster s.e.	0.23	**	999	0.39	***	999	-0.02		998	-0.08		999	1.36		994	0.06	*	994	-0.48		999	-0.01		999	-0.46		980
PSW Bootstrapped s.e.	0.23	**	999	0.39	***	999	-0.02		998	-0.08		999	1.36		994	0.06		994	-0.48		999	-0.01		999	-0.46		980

	Food Orientation																										
Methodology	Habit change perception on food <u>consumption</u> by Household (Categorical Index)			Habit change perception by Household (Continuous Index)			Habit change perception by Household (Categorical Index)			Habit change perception on food <u>selection</u> by Beneficiary (*Continuous Index*)			Habit change perception on food <u>selection</u> by Beneficiary (*Categorical Index*)			Habit change perception on food <u>consumption</u> by Beneficiary (*Continuous Index*)			Habit change perception on food <u>consumption</u> by Beneficiary (*Categorical Index*)			Habit change perception by Beneficiary (*Continuous Index*)			Habit change perception by Beneficiary (*Categorical Index*)		
	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N
PSM Kernel	-0.04		1010	0.041		1010	0.012		1010	0.508		1010	0.056		1010	-2.44		1010	-0.06		1010	-0.69		1010	-0.05		1010
PSM Nearest Neighbor	0.015		933	-0.85		930	-0.03		930	-0.72		936	0.029		936	-1.36		935	-0.03		935	-1.02		935	-0.06		935
PSM Stratification	0.022		1010	-0.3		1010	-0		1010	0.371		1010	0.057		1010	-2.49		1010	-0.06		1010	-0.8		1010	-0.05		1010
PSW robust & cluster s.e.	0.04		980	-0.06		975	-0		975	0.95		999	0.094	**	999	-3.65	*	984	-0.08	**	984	-1.03		984	-0.04		984
PSW Bootstrapped s.e.	0.04		980	-0.06		975	-0		975	0.95		999	0.094	**	999	-3.65	*	984	-0.08	**	984	-1.03		984	-0.04		984

	Food Orientation																		Health								
Methodology	Diet <u>Diversity</u> by Household (*Continuous Index*)			Diet <u>Diversity</u> by Household (*Continuous Index*)			Diet <u>Variety</u> by Household (*Continuous Index*)			Diet <u>Variety</u> by Household (*Continuous Index*)			Diet <u>Quality</u> by Household (*Continuous Index*)			Diet <u>Quality</u> by Household (*Continuous Index*)			<u>Diarrhea</u> symptoms in the Household (Ordinal Categorical Variable)			<u>Breathing</u> difficulties in the Household *(Ordinal Categorical Variable)*			<u>Yellowish skin</u> in the Household (*Binary Categorical variable*)		
	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N
PSM Kernel	-0.46		1010	-0.02		1010	-0.3		1010	-0		1010	-0.76		1010	-0.01	*	1010	-0.05		1010	0.037		1010	-0.02		1010
PSM Nearest Neighbor	-0.34		936	0.002		936	-0.24		936	-0.02		936	-0.59		936	-0.01		936	-0.03		936	0.033		936	-0.02		936
PSM Stratification	-0.31		1010	-0.02		1010	-0.26		1010	-0.02		1010	-0.57		1010	-0.01	*	1010	-0.05		1010	0.048		1010	-0.01		1010
PSW robust & cluster s.e.	-0.37		999	-0.02		999	-0.24		999	-0.72		999	-0.61		999	-0.01	*	999	-0.03		999	0.041		999	-0.01		999
PSW Bootstrapped s.e.	-0.37		999	-0.02		999	-0.24		999	-0.72		999	-0.61		999	-0.01	*	999	-0.03		999	0.041		999	-0.01		999

	Health																				
Methodology	<u>Eyes</u> disease symptoms in the Household (Binary Categorical *Variable*)			<u>Gum</u> disease symptoms in the Household (Binary Categorical *Variable*)			Beneficiary's <u>Diarrhea</u> symptoms (Ordinal Categorical Variable)			Beneficiary's <u>Breathing</u> difficulties (Ordinal Categorical Variable)			Beneficiary's <u>Yellowish skin</u> (Binary Categorical variable)			Beneficiary's <u>Eyes</u> disease symptoms (Binary Categorical Variable)			Beneficiary's <u>Gum</u> disease symptoms (Binary Categorical Variable)		
	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N
PSM Kernel	-0.03		1010	-0.03		1010	0.045		1010	0.026		1010	-0.02		1010	-0		1010	-0.01		1010
PSM Nearest Neighbor	-0.02		936	-0.03		936	0.006		936	0.013		936	-0.03	*	936	0.014		936	-0		936
PSM Stratification	-0.02		1010	-0.02		1010	0.044		1010	0.04		1010	-0.01		1010	-0		1010	-0.01		1010
PSW robust & cluster s.e.	-0.01		999	-0.01		999	0.05		999	0.02		999	-0.01		999	-0.01		999	-0		999
PSW Bootstrapped s.e.	-0.01		999	-0.01		999	0.05		999	0.02		999	-0.01		999	-0.01		999	-0		999

Note: * is significant at the 90% confidence level, ** is significant at the 95% and *** is significant at the 99%. Estimations with bootstrapped standard errors are replicated 100 times. Robust standard errors are clustered at the locality level. The causal effect is a simple difference, since the response variables are only captured in a single point of time. The propensity score estimations include municipality fixed effects, while the structural equations consider locality fixed effects.

INC Girls

	Food Support												Food Orientation														
Methodology	WAZ			HAZ			WHZ			BMI			Habit change perception on food <u>selection</u> by Household (Continuous Index)			Habit change perception on food <u>selection</u> by Household (Categorical Index)			Habit change perception on food <u>preparation</u> by Household (Continuous Index)			Habit change perception on food <u>preparation</u> by Household (Categorical Index)			Habit change perception on food <u>consumption</u> by Household (Continuous Index)		
	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N
PSM Kernel	0.314	***	940	0.36	**	940	0.21		940	0.143		940	1.074		940	-0		940	1.31		940	0.008		940	0.765		940
PSM Nearest Neighbor	0.386	***	884	0.301		884	0.36	**	884	0.314	*	884	0.528		882	0		882	0.572		884	-0.01		884	1.492		877
PSM Stratification	0.298	***	940	0.338	**	940	0.204		940	0.14		940	0.725		940	-0.01		940	0.884		940	0.001		940	0.444		940
PSW robust & cluster s.e.	0.39	***	943	0.48	***	943	0.207	*	941	0.13		943	0.83		940	-0.02		940	1.23		943	0.005		943	0.14		924
PSW Bootstrapped s.e.	0.39	***	943	0.48	***	943	0.207		941	0.13		943	0.83		940	-0.02		940	1.23		943	0.005		943	0.14		924

	Food Orientation																										
Methodology	Habit change perception on food <u>consumption</u> by Household (Categorical Index)			Habit change perception by Household (Continuous Index)			Habit change perception by Household (Categorical Index)			Habit change perception on food <u>selection</u> by Beneficiary (*Continuous Index*)			Habit change perception on food <u>selection</u> by Beneficiary (*Categorical Index*)			Habit change perception on food <u>consumption</u> by Beneficiary (*Continuous Index*)			Habit change perception on food <u>consumption</u> by Beneficiary (*Categorical Index*)			Habit change perception by Beneficiary (*Continuous Index*)			Habit change perception by Beneficiary (*Categorical Index*)		
	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N
PSM Kernel	0.037		940	1.156		940	0.096	**	940	3.69	**	940	0.133	**	940	2.293		940	0.016		940	3.64	**	940	0.077	*	940
PSM Nearest Neighbor	0.013		877	0.934		875	0.015		875	2.9	*	884	0.114	**	884	1.745		877	-0.01		877	2.83	*	877	0.036		877
PSM Stratification	0.035		940	0.687		940	0.076	**	940	3.5	**	940	0.124	**	940	1.295		940	-0		940	3.1	**	940	0.056		940
PSW robust & cluster s.e.	0.008		924	0.63		921	0.08	**	921	3.42	***	940	0.1	**	940	0.22		923	-0.02		923	2.38	**	920	0.05		920
PSW Bootstrapped s.e.	0.008		924	0.63		921	0.08	**	921	3.42	***	940	0.1	**	940	0.22		923	-0.02		923	2.38	*	920	0.05		920

	Food Orientation																		Health								
Methodology	Diet <u>Diversity</u> by Household (*Continuous Index*)			Diet <u>Diversity</u> by Household (*Continuous Index*)			Diet <u>Variety</u> by Household (*Continuous Index*)			Diet <u>Variety</u> by Household (*Continuous Index*)			Diet <u>Quality</u> by Household (*Continuous Index*)			Diet <u>Quality</u> by Household (*Continuous Index*)			<u>Diarrhea</u> symptoms in the Household (Ordinal Categorical Variable)			<u>Breathing</u> difficulties in the Household *(Ordinal Categorical Variable)*			<u>Yellowish skin</u> in the Household (*Binary Categorical variable*)		
	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N
PSM Kernel	-0.77		940	-0.02		940	-0.14		940	0.027		940	-0.91		940	-0		940	-0		940	-0.07		940	0.026	**	940
PSM Nearest Neighbor	-0.12		884	0.025		884	-0.44	**	884	-0.01		884	-0.56		884	0		884	-0.06		884	-0.05		884	0.037	***	884
PSM Stratification	-0.53		940	-0.01		940	-0.16		940	0.028		940	-0.69		940	0		940	-0.01		940	-0.07		940	0.026	*	940
PSW robust & cluster s.e.	-0.26		943	0.012		943	-0.14		943	0.02		943	-0.4		943	-0		943	-0		943	-0.1		943	0.03	**	943
PSW Bootstrapped s.e.	-0.26		943	0.012		943	-0.14		943	0.02		943	-0.4		943	-0		943	-0		943	-0.1		943	0.03	*	943

	Health																				
Methodology	<u>Eyes</u> disease symptoms in the Household (Binary Categorical *Variable*)			<u>Gum</u> disease symptoms in the Household (Binary Categorical *Variable*)			Beneficiary's <u>Diarrhea</u> symptoms (Ordinal Categorical Variable)			Beneficiary's <u>Breathing</u> difficulties (Ordinal Categorical Variable)			Beneficiary's <u>Yellowish skin</u> (Binary Categorical variable)			Beneficiary's <u>Eyes</u> disease symptoms (Binary Categorical Variable)			Beneficiary's <u>Gum</u> disease symptoms (Binary Categorical Variable)		
	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N
PSM Kernel	0.013		940	0.008		940	0.007		940	0.004		940	0.003		940	0.002		940	-0.01		940
PSM Nearest Neighbor	0.033		884	-0.01		884	-0.03		884	-0.03		884	0.009		884	-0.01		884	-0.04	**	884
PSM Stratification	0.018		940	0		940	-0		940	0		940	0		940	0.002		940	-0.02		940
PSW robust & cluster s.e.	0.02		943	-0.01		943	0.007		943	-0.02		943	0.004		942	0.01		943	-0.03	**	943
PSW Bootstrapped s.e.	0.02		943	-0.01		943	0.007		943	-0.02		943	0.004		942	0.01		943	-0.03	**	943

Note: * is significant at the 90% confidence level, ** is significant at the 95% and *** is significant at the 99%. Estimations with bootstrapped standard errors are replicated 100 times. Robust standard errors are clustered at the locality level. The causal effect is a simple difference, since the response variables are only captured in a single point of time. The propensity score estimations include municipality fixed effects, while the structural equations consider locality fixed effects.

INC Urban

	Food Support												Food Orientation														
Methodology	WAZ			HAZ			WHZ			BMI			Habit change perception on food selection by Household (Continuous Index)			Habit change perception on food selection by Household (Categorical Index)			Habit change perception on food preparation by Household (Continuous Index)			Habit change perception on food preparation by Household (Categorical Index)			Habit change perception on food consumption by Household (Continuous Index)		
	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N
PSM Kernel	**0.227**	*	**1443**	**0.435**	***	**1443**	-0		1443	-0.09		1443	1.286		1443	0.029		1443	0.736		1443	-0		1443	-0.38		1443
PSM Nearest Neighbor	**0.181**		**1357**	**0.285**		**1357**	0.054		1357	-0		1357	1.781		1352	0.039		1352	1.063		1357	-0.01		1357	-0.45		1357
PSM Stratification	**0.209**	*	**1443**	**0.453**	***	**1443**	-0.05		1443	-0.14		1443	1.139		1443	0.034		1443	0.461		1443	-0.01		1443	-0.75		1443
PSW robust & cluster s.e.	**0.21**	**	**1432**	**0.465**	***	**1432**	-0.07		1429	-0.15		1432	1.1		1424	0.035		1424	0.86		1432	-0		1432	-0.44		1404
PSW Bootstrapped s.e.	**0.21**	*	**1432**	**0.465**	**	**1432**	-0.07		1429	-0.15		1432	1.1		1424	0.035		1424	0.86		1432	-0		1429	-0.44		1404

	Food Orientation																										
Methodology	Habit change perception on food consumption by Household (Categorical Index)			Habit change perception by Household (Continuous Index)			Habit change perception by Household (Categorical Index)			Habit change perception on food selection by Beneficiary (*Continuous Index*)			Habit change perception on food selection by Beneficiary (*Categorical Index*)			Habit change perception on food consumption by Beneficiary (*Continuous Index*)			Habit change perception on food consumption by Beneficiary (*Categorical Index*)			Habit change perception by Beneficiary (*Continuous Index*)			Habit change perception by Beneficiary (*Categorical Index*)		
	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N
PSM Kernel	0.023		1443	0.392		1443	0.058		1443	1.241		1443	**0.078**		**1443**	-1.64		1443	-0.05		1443	0.164		1443	-0.02		1443
PSM Nearest Neighbor	0		1357	0.567		1344	0.043		1344	1.159		1357	**0.063**		**1357**	-1.92		1351	-0.03		1351	0.001		1351	-0.02		1351
PSM Stratification	0.013		1443	0.119		1443	0.045		1443	1.542		1443	**0.089**	*	**1443**	-2.3		1443	-0.07		1443	0.091		1443	-0.03		1443
PSW robust & cluster s.e.	0.01		1404	0.34		1396	0.05	**	1396	1.52		1429	**0.09**	**	**1429**	-2.65	*	1408	-0.07	**	1408	-0.15		1405	-0.03		1405
PSW Bootstrapped s.e.	0.01		1404	0.34		1396	0.05	*	1396	1.52		1429	**0.09**	**	**1429**	-2.65	*	1408	-0.01	**	1408	-0.15		1405	-0.03		1405

	Food Orientation																		Health								
Methodology	Diet Diversity by Household (*Continuous Index*)			Diet Diversity by Household (*Continuous Index*)			Diet Variety by Household (*Continuous Index*)			Diet Variety by Household (*Continuous Index*)			Diet Quality by Household (*Continuous Index*)			Diet Quality by Household (*Continuous Index*)			Diarrhea symptoms in the Household (Ordinal Categorical Variable)			Breathing difficulties in the Household *(Ordinal Categorical Variable*)			Yellowish skin in the Household (*Binary Categorical variable*)		
	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N
PSM Kernel	-0.49		1443	-0.01		1443	-0.05		1443	0.025		1443	-0.54		1443	**-0.01**	**	**1443**	-0.04		1443	-0.01		1443	0.009		1443
PSM Nearest Neighbor	-0.47		1357	0.003		1357	-0.11		1357	0.01		1357	-0.58		1357	**-0.01**		**1357**	-0.06		1357	-0.08		1357	0.009		1357
PSM Stratification	-0.31		1443	-0		1443	-0.02		1443	0.021		1443	-0.34		1443	**-0.01**	**	**1443**	-0.03		1443	-0.02		1443	0.011		1443
PSW robust & cluster s.e.	-0.14		1432	0.009		1432	-0.02		1432	0.01		1432	-0.16		1432	**-0.01**	***	**1432**	-0.01		1432	-0.01		1432	0.01		1432
PSW Bootstrapped s.e.	-0.14		1432	0.009		1432	-0.02		1432	0.01		1432	-0.16		1432	**-0.01**	**	**1432**	-0.01		1432	-0.01		1432	0.01		1432

	Health																				
Methodology	Eyes disease symptoms in the Household (Binary Categorical *Variable*)			Gum disease symptoms in the Household (Binary Categorical *Variable*)			Beneficiary's Diarrhea symptoms (Ordinal Categorical Variable)			Beneficiary's Breathing difficulties (Ordinal Categorical Variable)			Beneficiary's Yellowish skin (Binary Categorical variable)			Beneficiary's Eyes disease symptoms (Binary Categorical Variable)			Beneficiary's Gum disease symptoms (Binary Categorical Variable)		
	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N
PSM Kernel	-0.02		1443	-0.02		1443	-0.02		1443	0.027		1443	-0		1443	0.013		1443	-0.01		1443
PSM Nearest Neighbor	-0.01		1357	-0.02		1357	-0.05		1357	0.048		1357	0		1357	0.026		1357	-0.01		1357
PSM Stratification	-0.01		1443	-0.01		1443	-0.03		1443	0.045		1443	-0		1443	0.012		1443	-0.01		1443
PSW robust & cluster s.e.	-0.01		1432	-0.01		1432	0.006		1432	0.01		1432	0.003		1431	0.006		1432	-0.01		1432
PSW Bootstrapped s.e.	-0.01		1432	-0.01		1432	0.006		1432	0.01		1432	0.003		1431	0.006		1432	-0.01		1432

Note: * is significant at the 90% confidence level, ** is significant at the 95% and *** is significant at the 99%. Estimations with bootstrapped standard errors are replicated 100 times. Robust standard errors are clustered at the locality level. The causal effect is a simple difference, since the response variables are only captured in a single point of time. The propensity score estimations include municipality fixed effects, while the structural equations consider locality fixed effects.

INC Rural

	Food Support												Food Orientation														
Methodology	WAZ			HAZ			WHZ			BMI			Habit change perception on food selection by Household (Continuous Index)			Habit change perception on food selection by Household (Categorical Index)			Habit change perception on food preparation by Household (Continuous Index)			Habit change perception on food preparation by Household (Categorical Index)			Habit change perception on food consumption by Household (Continuous Index)		
	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N
PSM Kernel	**0.384**	***	**512**	0.155		512	**0.421**	*	**512**	**0.423**		**512**	1.82		512	0.035		512	0.351		512	0.017		512	2.023	*	512
PSM Nearest Neighbor	**0.395**	**	**474**	0.121		474	**0.473**	*	**474**	**0.483**		**474**	1.473		474	0.04		474	-0.55		474	-0		474	2.071		472
PSM Stratification	**0.367**	***	**512**	0.066		512	**0.47**	*	**512**	**0.485**	*	**512**	1.41		512	0.019		512	0.279		512	0.012		512	1.768	*	512
PSW robust & cluster s.e.	**0.56**	***	**510**	0.29		510	**0.557**	**	**510**	**0.566**	*	**510**	-0.09		510	-0.02		510	-1.27		510	0.001		510	1.16		500
PSW Bootstrapped s.e.	**0.56**	***	**510**	0.29		510	**0.557**	**	**510**	**0.566**	*	**510**	-0.09		510	-0.02		510	-1.27		510	0.001		510	1.16		500

	Food Orientation																										
Methodology	Habit change perception on food consumption by Household (Categorical Index)			Habit change perception by Household (Continuous Index)			Habit change perception by Household (Categorical Index)			Habit change perception on food selection by Beneficiary (*Continuous Index*)			Habit change perception on food selection by Beneficiary (*Categorical Index*)			Habit change perception on food consumption by Beneficiary (*Continuous Index*)			Habit change perception on food consumption by Beneficiary (*Categorical Index*)			Habit change perception by Beneficiary (*Continuous Index*)			Habit change perception by Beneficiary (*Categorical Index*)		
	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N
PSM Kernel	**0.104**	**	**512**	1.727		512	0.054		512	**4.471**	***	**512**	**0.127**	***	**512**	3.23		512	0.042		512	**4.51**	**	**512**	0.105	*	512
PSM Nearest Neighbor	**0.088**	*	**472**	1.242		472	0.015		472	**5.521**	***	**474**	**0.181**	***	**474**	3.587		472	0.012		472	**4.413**	***	**472**	0.101		472
PSM Stratification	**0.088**	**	**512**	1.453		512	0.045		512	**3.895**	**	**512**	**0.111**	**	**512**	4		512	0.065		512	**4.418**	**	**512**	0.101	*	512
PSW robust & cluster s.e.	**0.06**		**500**	-0.05		500	-0.03		500	**3.128**	**	**510**	**0.114**	***	**510**	2.03		499	0.017		499	**3.03**		**499**	0.06		499
PSW Bootstrapped s.e.	**0.06**		**500**	-0.05		500	-0.03		500	**3.128**	**	**510**	**0.114**	***	**510**	2.03		499	0.017		499	**3.03**	*	**499**	0.06		499

	Food Orientation																		Health								
Methodology	Diet Diversity by Household (*Continuous Index*)			Diet Diversity by Household (*Continuous Index*)			Diet Variety by Household (*Continuous Index*)			Diet Variety by Household (*Continuous Index*)			Diet Quality by Household (*Continuous Index*)			Diet Quality by Household (*Continuous Index*)			Diarrhea symptoms in the Household (Ordinal Categorical Variable)			Breathing difficulties in the Household *(Ordinal Categorical Variable*)			Yellowish skin in the Household (*Binary Categorical variable*)		
	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N
PSM Kernel	-1.18		512	-0.07		512	**-0.08**	***	**512**	**-0.09**	***	**512**	**-1.98**	**	**512**	-0		512	-0.09		512	-0.05		512	-0.01		512
PSM Nearest Neighbor	-1.42	*	474	-0.12	**	474	**-0.79**	***	**474**	**-0.1**	**	**474**	**-2.21**	***	**474**	-0		474	-0.12		474	-0.07		474	-0.01		474
PSM Stratification	-1.1		512	-0.06		512	**-0.77**	***	**512**	**-0.08**	**	**512**	**-1.87**	**	**512**	-0		512	-0.06		512	0.001		512	-0.01		512
PSW robust & cluster s.e.	-0.57		510	-0.02		510	**-0.82**	***	**510**	**-0.08**	**	**510**	**-1.39**		**510**	-0		510	0.05		510	-0.04		510	-0.02		510
PSW Bootstrapped s.e.	-0.57		510	-0.02		510	**-0.82**	***	**510**	**-0.08**	**	**510**	**-1.39**	*	**510**	-0		510	0.05		510	-0.04		510	-0.02		510

	Health																				
Methodology	Eyes disease symptoms in the Household (Binary Categorical *Variable*)			Gum disease symptoms in the Household (Binary Categorical *Variable*)			Beneficiary's Diarrhea symptoms (Ordinal Categorical Variable)			Beneficiary's Breathing difficulties (Ordinal Categorical Variable)			Beneficiary's Yellowish skin (Binary Categorical variable)			Beneficiary's Eyes disease symptoms (Binary Categorical Variable)			Beneficiary's Gum disease symptoms (Binary Categorical Variable)		
	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N	Coef.	Signif.	N
PSM Kernel	-0.01		512	0.007		512	0.123		512	-0.05		512	-0.02		512	-0.04		512	**-0.03**		**512**
PSM Nearest Neighbor	0.011		474	-0.09	**	474	0.126		474	-0.08		474	-0.03		474	-0.03		474	**-0.07**	*	**474**
PSM Stratification	-0.01		512	0.013		512	0.143		512	-0.03		512	-0.02		512	-0.03		512	**-0.03**		**512**
PSW robust & cluster s.e.	0.016		510	-0.02		510	0.127		510	-0.01		510	-0.01		510	-0.03		510	**-0.06**	*	**510**
PSW Bootstrapped s.e.	0.016		510	-0.02		510	0.127		510	-0.01		510	-0.01		510	-0.03		510	**-0.06**	**	**510**

Note: * is significant at the 90% confidence level, ** is significant at the 95% and *** is significant at the 99%. Estimations with bootstrapped standard errors are replicated 100 times. Robust standard errors are clustered at the locality level. The causal effect is a simple difference, since the response variables are only captured in a single point of time. The propensity score estimations include municipality fixed effects, while the structural equations consider locality fixed effects.