Ex-Ante Impact Evaluation: A Withdrawal of the Education Cash Transfer in Colombia

Daniel Flechas Buriticá

May 26, 2018

Abstract

I estimate an Intended-to-Treat (ITT) family behavioral response in terms of school attendance rate when hypothetically the educational cash transfer in Colombia withdraws (Todd and Wolpin, 2006). First, I use the theory of change from the Cash Transfer Program (CTP) to describe policy goals, results, activities and transmission mechanisms. Second, a family response model describes the CTP logical framework (Mantzavinos, 2006). Third, I provide the Potential Outcome Model (Rubin, 2005) in order to settling impact evaluation missing value problem and explore solutions for a counterfactual. Fourth, the Intended-to-Treat estimator is resolved using a bi-weight Kernel Regression to find the policy estimators for average school attendance rate. Two main data sets were used; on one hand from the CTP in Colombia, “Más Familias en Acción” (2016) I used school attendance information and, on the other, from the Household Survey (DANE, 2016) data about family income for beneficiaries and child labor wages as opportunity costs. I find an overall small negative effect with different variations according to socio-demographic dimensions age, gender, school grade, urban-rural. Lastly conclusions within the Institutional Analysis Development Framework (Ostrom, 2005) are drawn and discuss cost/benefit implications.

Keywords: Ex-ante Impact Evaluation, Public Policy, Cash Transfer Programs, Theory of Change, Educational Consumption Model, Kernel Density Functions, Intended-to-Treat.

*I would like to thank my Research Director Prof. Martha Misas Arango for her inspiring teachings in policy measurement and metrics.*
1. Cash Transfer Program Context

What if the education cash transfer is suspended? what is going to happen with the school attendance rate? how is going to be the response of beneficiaries? There is a causal mechanism that supports the logic of the Cash Transfer Program - CTP - useful for analytical purposes. An answer will have to provide information about the levels in outcome and impact indicators within an intuitive framework. This work provides a road map to build up results in scenarios of a single or multiple variables and provide estimations within its logic. I use an ex-ante evaluation strategy (Todd & Wolpin, 2006) on an policy situation (Ostrom, 2005) within the CTP in Colombia Mas Familias en Acción. The Program is a public policy that aims to work towards the reduction, prevention and mitigation of poverty and income inequity, to the formation of human capital and the strengthening of living conditions from poor and vulnerable families through a complementary subsidy to the income (BID, 2010; World Bank, 2009).

The theory of change (Binnendijk, 2000; Weiss 1995) describes the logic of operation within the public policy rationale (Figure No. A: Theory of Change: Educational Cash Transfer) and the impact in the Multidimensional Poverty Index (Figure B. Multidimensional Poverty Index). Program inputs are the local educational services and cash transfers (incentives); activities are mainly targeted to achieve the families decision to send their kids to school; outputs are the attendance certifications against which families receive cash transfers. Short term outcomes are a decrease in child labor and an increase in the school attendance rate. Those two sustained indicators over time attain medium term poverty outcomes measured by educational gap and educational achievement. In the long run, the program is reducing the illiteracy rate. Affecting those variables, the Program expects to impact negatively the poverty index. In Colombia, the Program Mas Familias en Accion’s objective is to reduce poverty levels by increasing human capabilities and conditions, the program has three main components, education, health and nutrition.

Over the last decades, the cash transfer model spread out rapidly throughout the world (Figure C: Conditional Cash Transfer Programs in the World: 1997 and 2008). According to CEPAL, by 2010 the cash transfer programs have been implemented in about 18 countries and have reached about 25 million families (about 113 million persons) or the 19% of the regional population at a cost of 0.4% of the Regional Gross Product (Cechini & Madariaga, 2011). In Colombia, according to public data in 2016, about 2.5 million families were enrolled in the program and they were benefiting about 3.14 million children1. Because of Program scaling-up; decisions and calculations are not easily at hand. Planning decision-making scenarios

---

1 http://www.prosperidadsocial.gov.co/Paginas/M%C3%A1s-Familias-en-Acc%C3%B3n.aspx retrieved May, 2016
require several calculations that imply huge amount of information barely out of the human brain’s capacity (Kahneman, 2011; Marois and Ivanoff, 2005); despite its impressive complexity and processing power. Policy decision makers are subject to this short-sighting constraints; therefore a problem of measurement attached to economic theory. Policy relevant estimators provide a good answer and tool to policy makers; they are simple, intuitive and concrete while carrying information from the economic model. They condense a vast amount of data while keep the scientific rigourity allowing high confidence on statistical populations.

The advance of CTP on the world over the last two decades brought various impact evaluations. The research agenda approaches ex post and ex ante studies. Some of the most relevant ex post impact evaluations in the region were developed by the International Food Policy Research Institute (Schultz, 2000, 2001; Sloufias, 2005) in Mexico with the program Oportunidades and in Nicaragua with the Red de Protección Social (Maluccio and Flores, 2004). In Ecuador, impact evaluations on the Bono de Desarrollo Humano (Oosterbeck, Ponce and Shady, 2008); in Jamaica the Programme of Advancement through Health and Education is evaluated by Levy and Ohls (2010). Relevant meta analysis on impact evaluations on Conditional CTP, specifically on school attendance rate and desertion done by Saavedra and García (2012). On the other hand, there are several researchers that choose for an ex-ante strategy; the most relevant are Bourguignon, Ferreira and Leite (2002, 2003) on the Bolsa Escola Program in Brazil, and the Oportunidades in Mexico (Todd and Wolpin, 2006, 2008). In Colombia, the main impact evaluations are lead by Attanasio and Gómez (2004; 2005; 2012) on long term effects in municipalities under 100 thousand inhabitants. Analysis on time series discuss the effectiveness on the program (Avila, 2012) and others measure the effects on human capital (Baez and Camacho, 2011).

In Latin America the overall results show a positive effect on school attendance rate and in human capital levels in families under poverty conditions (Rawlings and Rubio, 2005). Despite such evidence several remarks remain on the air: first, the effectiveness of the cash transfers when environmental changes among countries; second, the sustainability of social benefits; third, the scaling-up capacity to other poverty dimensions; and fourth, the effectiveness in intergenerational poverty transmission. Lastly, Ravallion (2005) states that there is no one single methodology for impact evaluations; instead, they should have an open-mind attitude and encourage future research agenda to dedicate resources to design relevant public policy analysis that get deeper understanding of the phenomena and enrich the results and conclusions on the outcomes for policy-makers and researchers.

---

2The human brain is bearded for its staggering complexity and processing capacity: its hundred billion (10/11) neurons and several hundred trillion synaptic connections can process and exchange prodigious amounts of information over a distributed neural network in the matter of milliseconds (Maraos and Ivanoff, 2005).

3"... clarity of thought is still a pearl of great price. In particular, the multiplicity of values of verbal symbols may be a great disadvantage when it comes to drawing the logical consequences of a proposition... " (Arrow, 1951a)
I am interested in predict the school attendance rate in a hypothetical case of a withdraw of Más Familias en Acción and support medium and long term analysis on the causal effects on outcome variables. I use an ex ante evaluation approach. It differs from an ex post impact evaluation in three dimensions. The first, is on its timing, an ex ante evaluation is done before the policy decision is made and therefore is a policy prediction. The second is on data availability in the sense that variables haven’t been realized yet in an ex ante approach versus an already observable realized variable in an ex post one. Under an ex ante evaluation only baselines are needed, in ex post evaluations at least two observations in different times in the process are requested. The third difference is the methodological design. While in the ex ante evaluation the treatment and control groups are hypothetical (since no real treatment happened yet), in the ex post evaluation treatment and control groups are real populations from which the experiment took place and therefore data will show the real effects in those differentiated populations.

I tackle the CTP withdraw question situating the decision-maker within an Institutional Analysis and Development Framework - IADF (North, 1990, 1991, 2005; Ostrom, 2005) as a game setter. With the use of statistics, build a counterfactual from an initial population and modeling with the use of economic theory. In doing so, the methodological road map (Judge et al., 1988) describes the research steps to answer the question. Based on the Program’s theory of change, a description of the real world phenomena and a causal-effect mechanism on how to impact some structural variables with short term behavioral models. Through a process of formalization and using the rules of logic (Mantzavinos, 2006) simulate a family’s choice model for consumption to estimate the behavioral response (Todd and Wolpin, 2008) in terms of school attendance rate. The answer is embedded in an IADF (Ostrom, 2005) as a tool to tell the story with a focus on an Action Arena: Más Familias en Acción. There are two main results: (a) descriptive - driving intuition; (b) empirical results - policy estimators. On the first, the model for a policy change shows IADF efficiency since keeps researcher intuition well situated among many dimensions and levels of reality. The analytical framework allow for a scientific approach to the educational component of Más Familias en Acción. The second, empirical results, show a heterogeneous effects and generally minimum negative impacts accordingly with theoretical statements.

2. FOUNDATIONS AND IDENTIFICATION STRATEGY

The Road Map: Methodological Guidance’s Choice

The configuration of the public policy model within an action arena integrates two ap-
proaches: postulation and experimental approaches (Judge et al., 1988). It contains an explicit behavioral causal model such that an empirical analysis is feasible. A big part of the knowledge process is derived from an abstraction (Arrow, 1954a); thanks to them, economical, political or mathematical environments are human devised systems helpful for explanatory purposes. Judge’s approach to the scientific research starts from the observation and comprehension of the real world phenomena to be translated into propositions and use postulates. This is known as the Postulation Approach. In this work, the approach uses IADF to situate an Action Arena to understand the real world phenomena. I then use the rules of logic to come to theoretical conclusions driving to empirical questions, here logic provides an intuition framework describing the causation mechanism. The logic is captured in a model which represents a simplified version of a multifaceted reality that is way bigger. It provides a description of variables and mechanisms that explain partially the social phenomena. The postulates that constitute the abstraction process derive from deductive logic inferences on propositions (Judge et. al., 1988, p.4). The Experimental Approach uses the conclusions from the postulation approach and test the real-world truth content. A theoretical model is proposed and the elements are handled by the use of probability to reach statistical conclusions to be compared with the postulation propositions. Given these knowledge-search processes, the task of econometrics is to use postulation and experimentation to generate information useful to reach descriptive and prescriptive conclusions about economic processes and institutions (Figure D: Research Methodology).

Real-World Phenomena: Setting the Game

The context determines the action arena and is set by configurations of physical conditions, rules-in-use, and the attributes of a community. The core of the action arena contains a set of variables, such as: stakeholders, their possible actions, positions, incentives, information, and cost-benefit analysis. The action arena allows an interaction among the participants given the context (Ostrom, 2005, 2010). This is the game setter (Figure E: Institutional Analysis and Development Framework). Easton (1953, 1957, 1965) defines public policy as the authoritative allocation of values for a society, so they are tools that public servants use to govern or regulate the action of the State. Policies have two central aspects: (i) plans that lead to specific goals, therefore an intention; (ii) a set of de facto decisions about allocation of resources aligned with such ends. As governing tools, public policies carry on action plans within; therefore the study of public policy implies the analysis of the stakeholder’s behavior (i.e., local authorities, citizens, political parties, private companies, etc.) (Ostrom, 2005) that might follow some patterns (Goodin, Morin & Rein, 2006). Policy measurement techniques are oriented to understand,

---

4Decision theory is concerned with the reasoning underlying an agent’s choices, whether this is a mundane choice between taking the bus or getting a taxi, or a more far-reaching choice about whether to pursue a demanding political career. (Note that “agent” here stands for an entity, usually an individual person, that is capable of deliberation and action) Standard thinking is that what an agent does on any given occasion is completely determined by her beliefs and desires/values. The orthodox normative decision theory, expected utility (EU) theory, essentially says that, in situations of uncertainty, one should prefer the option with greatest expected desirability or value (Steele & Stefansson, 2015)
structure, capture, estimate and communicate the elements and transmission mechanisms inside such “black box” (Easton, 1957).

Public policies might be nested in four levels (Figure F: Levels of Analysis): meta-constitutional situations (L1), constitutional situations (L2), collective-choice situations (L3), operational situations (L4) (Ostrom, 2005; Williamson, 2000, 2009). The upper levels impose restrictions over the rest; while the lower provide feedback to the superior levels in the form of information used in a change, adjustment or adaptation. The meta-constitutional situation (L1) refers to those where informal institutions, culture, traditions, norms and beliefs are located. In most cases, public policies take this information as exogenous and are the typical variables that provide the environment. Institutions at this level have a very low change ratio, many times take generations or centuries. Changes happen in the long term and windows of opportunity for changes are not very often. Most of the norms are not processed by individuals and are spontaneous responses. Given its evolutive origin, these norms are just assumed by agents and impact the inertia of decisions. The constitutional situations (L2) are structured from evolutive processes and human design. At this level formal institutions (North, 1990) appear in the form of constitutions, laws, decrees. These norms are shaped and constrained by the meta-constitutional level and often describe the type of governance structure a society has (for instance a democratic and create the three branches), as well as the bureaucratic rules, property rights, etc. The third level is the collective choice situation (L3) where the institutions of governance and public policies are located. Public policies are the societal devices to govern communities and translate rights into contracts among citizens. The fourth level is the operational situation (L4) where the daily transactions take place and implementation of policies factually happen. At this level, agents make daily decisions on their best allocation of resources, production or consumption. Operational situations tend to change in a permanent basis (Williamson, 2000).

Looking at the four levels of analysis mentioned above, the school attendance rate is an operational situation decision-type linked to a cash transfer program and eventually allow for the effects at the public-choice level (medium and long term effects as outcome and impact indicators respectively). The program’s design and implementation is at the collective choice level. The program responds to a social policy objective of poverty reduction that uses the public policy as a function to transform the poverty levels. The Program is linked to the society’s goal; which is set at the constitutional level. The Program is one of the largest social programs (about 150 million USD in 2016 for the educational component5) targeted to vulnerable population funded with national budget and, earlier on public international credits. The Program started in 2001 as a governmental response to the economic crisis in the decade of the 90’s (DNP, 2008). Más Familias en Acción has three main components, education, health and nutrition. At the collective-decision level, the Public Policy seeks to reduce poverty indices through a

---

5www.minsulcienago.gov.co
multi-annual program. At the operational level the challenge is to allocate cash transfers to elected households conditioned to a set of requisites. Household are the target population (unit of treatment) and have to meet the following requirements: (i) a poverty condition according to a marginality index; (ii) have children under 18 years old; (iii) application before the municipal authority.

The key stakeholders at this level are the Presidential Program (responsible for CTP overall implementation); Minister of Education (design and reinforce a national public education policy, oversees regional education plans that guarantee facilities, teachers and administrative staff) and banking entities (making the cash transfer from government to families). Players at this level can affect the parameters of the game (i.e. thresholds for eligibility, focusing policies, budgetary decisions, education supply). The cost/benefit ratio will measure the cost of the Program against the Multidimensional Poverty Index. The key stakeholders at local levels are municipal liaisons (Program officers responsible for identifying eligible families, help them in their application forms and, once in the program, guarantee compliance as well as the public institutions), the public schools (provide public education in a regular basis and following the educational standards) and the head of household (who receives the cash transfers and make the decision about child attendance to school or send them to work). The incentives are delivered through education cash transfers provided to children between 5 and 18 years old that are enrolled in the public education system from kindergarten to 11th grade. The incentives are conditioned to at least 80% of their monthly attendance to classes and up to 2 years in the same grade. Identification is based on the program’s assumptions; (i) budget constraints impede household head to send the child to school because of its direct costs and opportunity costs; (ii) the low levels in social capital as a fundamental reason for inter-generational poverty; (iii) well designed incentives can promote positive behavioral changes. The pattern of interaction is represented by the family’s behavioral model that simulates the decision-making mechanism responding to a cash transfer withdrawal at the operational level (Williamson, 2000; Ostrom, 2005). (Figure G: Action Arena: Familias en Acción).

**Tractable Reality**

Propositional logic as a tool allows to use a set of premises (made out of propositions) from real-world phenomena. Public policy as social phenomena is susceptible to be described through propositions and logic connectors. According to Arrow (1951) the language is the communication channel between the phenomena and the abstraction process. The use of formal language along with qualitative descriptions offers rigourosity, complexity and dynamism. Qualitative descriptions offer learning opportunities, intuition on the institutional framework, the stakeholders characteristics, the processes and the causal representation. The axiomatization of reality in theoretical systems with the use of models offers an alternative for a description of social phenomena. Mantzavinos (2007) proposes the use of propositional logic as a technique
to build a scientific explanation of a social fact. Deductive arguments take a set of initial conditions and a set of laws that govern the combination produces an effect, that can be a human action. The proposition about the state of the world to be explained (explanandum), is logically derived from the combination of singular propositions (initial conditions - C) and general laws (explanans - L) (Mantzavinos, 2007). With this logical structure one can build up a hypothesis of the type: consider a set of initial conditions with rationality as the behavioral law that produce an action (s). The conclusion (s), is a deduction of union of the propositions C and L.

Let's assume a household makes the decision to send the kid to school (which becomes the output variable of the social function s) taking into account the consumption level c. The restriction depends on the household income y, the opportunity cost of child labor w and the incentive τ to school attendance; which is designed to promote a certain behavior. A natural withdraw means the absence of the incentive at the family level. According to theory of change, families might use the opportunity cost in child labor as a compensation mechanism to cover the gap left behind. According to Mantzavinos’ propositional approach, a set of initial conditions can be picked, say (y, w, s, τ, X); while families make rational decisions governed behavioral law L, which produce an observable outcome in s; as shown in Figure No. 1.

With this logical structure one can build up a hypothesis of the type: consider a set of initial conditions (C) with rationality (L) as the behavioral law that produce an action (s) (Mantzavinos, 2007). This can be written as:

\[ y \land w \land \tau \land X \land L \rightarrow s \]

**Family Model for Schooling**

Supported in the theory of rational choice\(^6\) Todd and Wolpin (2006, 2008) developed a household decision-making model for Cash Transfer Programs regarding the education component. The Todd-Wolpin model was designed for a program’s entry. In this work, I focus on the exit following the same model. The household decides whether to send the child to school or to work. The household’s utility depend on the consumption (c) and on whether the child

\(^6\)See Appendix A: Rationality
attends to school (indicator \(s\)). According to the Program a child that doesn’t attend school is assumed to go for work in the labor market at a wage \((w)\). Letting \((y)\) denote household income, net of the child’s earnings, the household solves the problem of utility maximization with respect to school attendance \(\max_{\{s\}} U(c, s, \gamma)\) subject to \(c = y + w(1 - s)\). The optimal choice is \(s^* = \varphi(y, w, \gamma)\), where \(\varphi\) is a distribution function in terms of the family earnings, the wage in the labor market and \(\gamma\), an unobservable heterogeneity. Under the Program, the household receives an incentive for school attendance \((\tau)\) and \(\bar{y} = y + \tau\) represents the income positively shocked and \(\bar{w} = w - \tau\); which represents the child’s time taken from the labor market and represents the opportunity cost. Families that belong to the program will be those with income \(\bar{y}\) and child wage \(\bar{w}\). The maximization problem now changes the constraint set, so that \(\max_{\{s\}} U(c, s, \gamma)\) subject to \(c = (y + \tau) + (w - \tau)(1 - s)\) or \(c = \bar{y} + \bar{w}(1 - s)\). The optimal choice is now \(s^{**} = \varphi(\bar{y}, \bar{w}, \gamma)\).

I follow here the Todd and Wolpin (2008) model in a hypothetical scenario of a program withdrawal: incentive cancellation as a negative shock. With the exit of the program, population’s income is \(\bar{y} - \tau = y\) and using the program model the family might compensate the loss in the labor market \(\bar{w} + \tau = w\). The family will resolve this problem:

\[
\begin{align*}
\max_{\{s\}} U(c, s, \gamma) \\
\text{st.} \\
\qquad c = y + w(1 - s)
\end{align*}
\]

the optimal choice of \(s\) in the absence of the subsidy is \(s^{***} = \varphi(y, w, \tau, \gamma)\), where \(y = \bar{y} - \tau\) represents the income negatively shocked and \(w = \bar{w} + \tau\), the child’s time dedicated to the labor market and indicates the opportunity cost or the compensation reaction. Treated families \(D = 1\) are those that have the program suspended with restriction levels at \(c = (\bar{y} - \tau) + (\bar{w} + \tau)(1 - s)\) and control families \(D = 0\) they who belong to the program and restriction levels at \(c = \bar{y} + \bar{w}(1 - s)\).

**Figure 2: Program Determinants in Three Stages**

<table>
<thead>
<tr>
<th>household income</th>
<th>Absence of Program</th>
<th>Program Implementation</th>
<th>Program Withdrawal</th>
</tr>
</thead>
<tbody>
<tr>
<td>child labor</td>
<td>(y)</td>
<td>(\bar{y} = y + \tau)</td>
<td>(y = \bar{y} - \tau)</td>
</tr>
<tr>
<td>subsidy</td>
<td>(w)</td>
<td>(\bar{w} = w - \tau)</td>
<td>(w = \bar{w} + \tau)</td>
</tr>
<tr>
<td>heterogeneity shock</td>
<td>(\gamma)</td>
<td>(\gamma)</td>
<td>(\gamma)</td>
</tr>
<tr>
<td>distribution function</td>
<td>(\varphi(y, w, \gamma))</td>
<td>(\varphi(\bar{y}, \bar{w}, \gamma))</td>
<td>(\varphi(y, w, \tau, \gamma))</td>
</tr>
<tr>
<td>optimal choice</td>
<td>(s^* = \varphi(y, w, \gamma))</td>
<td>(s^{**} = \varphi(\bar{y}, \bar{w}, \gamma))</td>
<td>(s^{***} = \varphi(y, w, \tau, \gamma))</td>
</tr>
</tbody>
</table>

Now, the evaluation question is: What is the change in probability of family \(i\) to send its kid to school under an exit of the Program?
Potential Outcome Model and Policy Estimator

The challenge that arises to answer this question is known as the fundamental problem of causal inference and one way to come around is using the Potential Outcome Model (Holland, 1986; Rubin, 1974, 2005); which helps to find counterfactual within data. The problem of causal inference refers to a measurement and inference that explains a causal link between the treatment and the result (Holland, 1986). The impact evaluation implies the effect of a variable of interest; and such effect refers to a change in the socioeconomic status of the agents. An observed effect (s) with a treatment (D) to households with characteristics (X), configure the set of information (s, X, D). Over this set, it is possible to estimate a change in s, conditioning for D, holding X constant. The treatment D is a categorical variable that takes the value of 1 when treated and 0 when non-treated. Given that an agent can be either treated or non-treated but not both at the same time, then the fundamental problem arises with mutually exclusive events. In this case, only one observation is available and the missing value is the counterfactual. A counterfactual is just hypothetical and need to be contrary to certain facts (Heckman, 2005). The effect of the treatment on a family is measured by (s_1 – s_0); one observation and its counterfactual. This first estimation requires a comparison between the set of treated elements and non-treated elements (Cameron & Trivedi, 2005 p. 34). Since each unit receives only one treatment, either s_0 or s_1 is observed but not both (is missing for one unit), so comparisons of s_0 and s_1 imply some degree of speculation to a sense (Rubin, 1983). The assumption that there is a unique value s_{1i} corresponding to unit s_{0i} or i and treatment τ is called Stable-Unit Value Assumption SUTVA (Rosenbaum & Rubin, 1983). An Average Treatment Effect (ATE) is one way to resolve it (Rubin, 1974, 2005). At this point, the experimental approach requires a randomization between the treated and non-treated, used as controls. A randomize selection guarantees that treatments are not correlated with some population’s features. Now, N units viewed as a single random sample from some population, and the quantity to be estimated is the average between the treated and non-treated. Given the missing data problem, the expected value is built on the probability distribution F(.) from population.

\[ ATE = E[s \mid D = 1] - E[s \mid D = 0] \]

Contrary to an ex post evaluation, the ex ante faces the problem of missing information at the treatment level, so the expected value of the school attendance rate can’t be estimated. Therefore the ATE estimator is not useful for an ex ante methodology and a new estimator will be needed. Following Todd and Wolpin (2008) ex ante estimation of effects, I compare children from treated families with income \( y \) and child wage \( w \) and non-treated families with income \( \bar{y} \) and child wage offers \( \bar{w} \). The comparability comes from the fact that both populations share the distribution function from the heterogeneity shock. The required assumption on the distribution of observed heterogeneity is
\[
    f(\gamma \mid y, w, x) = f(\gamma \mid \tilde{y}, \tilde{w}, x)
\]

A typical matching estimator (e.g., Rosenbaum and Rubin, 1983) would assume that there exists a set of observable \( X \) for family \( i \) such that outcomes are independent on program participation conditional on \( X \):

\[
    (s_{1i}, s_{0i}) \perp \!\!\! \perp D_i \mid X_i
\]

That is, the school attendance rate for family \( i \) is statistically independent from the selection process \( (D_i) \) given the observable set. Here the conventional matching approach is not useful for ex ante evaluation, because it requires data on \( s_i \) which is not observed. Todd and Wolpin (2008) provide a modified version of matching using the fact that the economic model along with the restriction on the distribution of \( \tau \) implies:

\[
    s_{1i} = s_{0j} \mid y_i = \tilde{y}_j - \tau, w_i = \tilde{w}_j + \tau
\]

Family \( i \) under exit program (treatment) will suffer a reduction in its income related to the current (control) family \( j \), that is \( (y_i < \tilde{y}_j) \). According to the behavioral model, the household will be forced to compensate its loss in the labor market by increasing its participation by \( \tau \), and making \( w_k > \tilde{w}_i \).

Now, the evaluation question is: What is the estimated probability of the family \( i \) to send their children to school using information of realized family \( j \)?

\[
    \text{Prob} \{ s_i = 1 \mid y_i = \tilde{y}_j - \tau, w_i = \tilde{w}_j + \tau \} = \text{Prob} \{ s_i = 0 \mid y_i = \tilde{y}_j - \tau, w_i = \tilde{w}_j + \tau \}
\]

\[
    = E(s_i \mid y_i = \tilde{y}_j - \tau, w_i = \tilde{w}_j + \tau)
\]

The expected value of the school attendance rate \( E(s_i) \) for the family \( i \) under withdrawal of the program is evaluated in two points \((\tilde{y}_i, \tilde{w}_i)\). Todd and Wolpin use this value to build the matching estimator Intended-to-Treat (ITT) for a subsidy level of \( \tau \) calculating the average on the population. An Intended-to-Treat (ITT) estimator is used to explore policy relevant effects of hypothetical situation in case of the cancellation (Angrist & Imbens, 1994; Todd & Wolpin, 1997, 2008; Heckman, 2000, 2001). Then, to resolve the expected value equation the authors use a biweight kernel density (see Appendix B) that provides an estimated density function (EDF) for the school attendance rate \( s_i \) for family \( i \). The density function is built from information on the data available and takes a sample from a region of support \( S_p = \{(y, w) \text{ such that } f_{y,w}(y, w) > 0\} \) and \( f(y, w) \) is the density from which counterfactuals are built. Thus, I compare treated households based on matching as method of evaluation. The kernel method constructs matches using all individuals in the comparison sample and down weighting distant observations (Heckman, Ichimura, Todd, 1998). Then evaluate the differences between treated and non-treated aggregated into a policy estimator ITT. Thus the ITT overcomes the problem of the missing value at the treatment level.
The propensity score is the conditional probability of assignment to a particular treatment given a vector of observed covariates. Comparisons come from \( s_{0i} \) and \( s_{1i} \), for instance \( s_{1i} - s_{0i} \) or \( \frac{s_{1i}}{s_{0i}} \) (Rubin & Rosenbaum, 1983). Todd & Wolpin suggest to follow the stochastic solution when base their statistical inference on Rubin’s work. Here, the Bayes Theorem is highly useful since I don’t know the probability of sending the kid to school or not. For grounding reasons (theory of change) I obtain densities for \( s \) based on prices variability, family income and child wages \((y, w)\). This solution allows for the experiment, since I will induce a monetary shock that will serve as proxy for probabilities of going to school given a set of conditions. This is equal to the probability of \( s \) when she goes to school times her probability to have those conditions given the kids went to school. This scenario is a restricted space against a non restricted space and allow for looking at marginal changes.

\[
Prob (s = 1 \mid X) = \frac{\text{prob} (s = 1) \text{prob} (X \mid s = 1)}{\text{prob} (X)}
\]

If I have a partition, then one can derive the total probability based on the different partitions aggregation process. The solution of total probability allow Todd and Wolpin use the Kernel because they work over sample space partitions. The denominator is the law of total probability, which states that a sample space can be subdivided in a countably finite partition and each event is measurable.

\[
Prob (X) = \text{prob} (s = 1) \text{prob} (X \mid s = 1) + \text{prob} (s = 0) \text{prob} (X \mid s = 0)
\]

It can be modeled using an appropriate logit model or discriminant score:

\[
Prob (s = 1 \mid X) = \frac{\text{prob} (s = 1) \text{prob} (X \mid s = 1)}{\text{prob} (s = 1) \text{prob} (X \mid s = 1) + \text{prob} (s = 0) \text{prob} (X \mid s = 0)}
\]

Given \( X = (\tilde{y}_i - \tau) + (\tilde{w}_i + \tau) \) a distributional assumption on log-wages is made since its its positive smoothing effects allow for bigger support region. Assume the expected value has a lognormal distribution \( \xi \), with parameters from Household Survey - DANE, 2016. Given the constraint \( c = y + w(1 - s) \), we have the ln wage offer equation (Todd and Wolpin, 2010):

\[
\ln (w) = \mu_w + \varepsilon
\]

The family chooses to send their child to school \((s = 1)\) if they follow the decision rule:

\[
U (w, s = 1) > U (w + \exp (\mu_w) \exp (\varepsilon), s = 0)
\]

On the ln wage distributional assumption assume that \( \varepsilon \) is normally distributed with mean 0 and variance equal to \( \sigma_{\varepsilon}^2 \) and that \( \varepsilon \) is distributed independently of family income \( f (\varepsilon \mid w) = f (\varepsilon) \). To take into account selectivity in observed wages, authors wrote the wage equation as

\[
\ln (w) = \mu_w + \mathbb{E} (\varepsilon \mid U (w + \exp (\mu_w) + \exp (\varepsilon), s = 0) > U (w, s = 1)) + \nu
\]
\[ \eta = \mu_w + E(\varepsilon \mid \varepsilon > \eta(w)) + \nu \]

where \( \nu \) has conditional mean zero by construction and \( \eta \) is some function of \( w \). The conditional mean function can be written as

\[ E(\varepsilon \mid \varepsilon > \eta(w)) = \frac{\int_{\eta(w)}^{\infty} \varepsilon f(\varepsilon) d\varepsilon}{\int_{\eta(w)}^{\infty} f(\varepsilon) d\varepsilon} \]

The question then is about the effect in \( \nu \); resolved estimating the matched outcomes \( E(\cdot) \) nonparametrically using a two dimensional kernel regression estimator, letting the hypothetical family \( y_0 = \bar{y}_j - \tau \) and \( w_0 = \bar{w}_j + \tau \) on family income and child wage, such as:

\[ E(s_i = 1 \mid y_i = y_0, w_i = w_0) = \frac{\sum_{i=1}^{n} s_j K \left( \frac{w_i - w_0}{h} \right) K \left( \frac{w_i - w_0}{h} \right)}{\sum_{i=1}^{n} K \left( \frac{w_i - w_0}{h} \right) K \left( \frac{w_i - w_0}{h} \right)} \]

and \( K(\cdot) \) denotes the biweight kernel function (see Appendix B) and \( h \) are the bandwidth parameters (Hansen, 2009). Having an expected value for one of the missing data one can build the ITT. The estimator provides the average values over the population between counterfactuals and observations. It resolves the expected value for the population at stake. The first term of the equation provides the expected value for the single family \( i \) (counterfactual) and discounts with the observed value of family \( j \) (baseline). The average can only be taken over the region of overlapping support \( S_p \), which in this case is over the set of families \( i \) for which the shocked values \( \bar{y}_j - \tau \) and \( \bar{w}_j + \tau \) lie within the observed support region.

\[ \text{ITT} = \hat{\alpha} = \frac{1}{n} \sum_{i=1}^{n} \{ E(s_i = 1 \mid y_i = y_0, w_i = w_0) - s_j (\bar{y}_j, \bar{w}_j) \} \]

ITT gathers the analytical outcomes of the action arena estimated at the population level according to key policy decision factors, i.e., age, school grade, gender, geographic location, incentive level (\( \hat{\alpha}_{\text{age}}, \hat{\alpha}_{\text{gender}}, \ldots \)). The expected value could include also other relevant sociodemographic conditions in order to gain accuracy and precision in the comparison process among families.

Concluding the conceptual framework, what is new in this paper is the pragmatic way to see foundations in economic theory in order to resolve an empirical question on a public policy field. It provides the analytical roots to build a formal scenario and measure their multivariate effects. The policy treatment effects are the quantitative results out of the behavioral model under the framing of the Action Arena; its elements help to “tell the story” about the cause-effect relationship. A methodological connection between estimators and action arenas provide a wider vision of the real-world phenomena. It also helps to see qualitative and quantitative
scenarios of public policy decision. The model estimates the hypothetical effects that have even deep consequences on impact variables. It might invoke some thought at higher levels such governance structures since intuition on scalar movements are better understood.

3. Data Section: Initial Conditions (C)

Two main data sets were used. The first from the CTP in Colombia, “Más Familias en Acción”, specifically the period March-April 2016 with about 3.36 million female and male kids from 5 to 18 years. The data set has sociodemographic information about the kids and has no information at family level so assume households have only one kid. The school attendance rate (s) provides information about accomplishment, that is a dichotomous variables with 1 if the kid attended to at least the 80% of the time for a two months period, and 0 otherwise. According to school grade and municipality category, she receives a fix amount of money each period (τ). The incentive levels are adjusted every year. The following four tables show the incentives participation by age, grade, group and location. The program has 10 incentive levels from zero to maximum at 62,475 (prices of 2016).

![Figure 3: Cash Transfer Table by School Grade and Group](image)

The population’s concentration by age takes the following form. Figure 2: Population Density, Age by Incentive Level. The population between 5 to 12 years old and incentive levels until the third level represent the 51% of the population. The peak are kids of ten years that receive the second incentive level (17,075 COP).

The opportunity of technology is to capture all information in datasets with easy and real time calculations, estimations, read with graphical data to visualize effects. The findings will be relevant for policy makers, public officers, beneficiaries, stakeholders.
Table No. 2, shows the population distribution by school grade and incentive levels. The kindergarten and primary education are getting the lowest levels of incentives until the third (22,725 COP) and represent the 56% of the population.

Figure 5: Population Density, School Grade by Incentive Level

Data Integration

From the Great Integrated Household survey -GEIH- (DANE, 2016-IV), include the Child Labor Module. From here, identify the household population that belongs to Más Familias en Acción and build the aggregated household income ($y$) and the child labor wage ($w$). I estimate household income through the aggregation household income according to the methodology provided by Mision para el Empulme de las Series de Empleo, Pobreza y Desigualdad (MESEP) (DNP-DANE, 2012). The aggregated household income comes from the population over 12 years old and the economically active. There exist four categories of income gatherers: (i) salaried employee, (ii) independent, (iii) family workers without regular remuneration, and (iv) unemployed or inactive. The income receiver can apply for the different types of income: (i) monetary income first activity; (ii) income in-kind; (iii) income secondary activity; (iv) monetary income of inactive or unemployed; and (v) income from other sources. The household income ($y$) is the aggregation or receivers’ income flow and represented by the indicator available.
income for consume per household. I used the child labor module (between 5 and 17 years old) from the GEIH to aggregate the child labor wage \((w)\). The aggregation uses the same categories created as an income receiver. Both values, household income and child labor opportunity cost are assigned to every kid. I use a truncated lognormal distribution conditioned by department to provide household income \((y)\). For the child labor values I used also a truncated lognormal conditioned by rural/urban, age and sex \((w)\). The distributional values come from the GEIH dataset. As shown in Figure 6: level versus lognormal distributions, monetary variables are going to be treated in logarithm given the smoothing effects and therefore increase in the support region as the following graph shows. The heat map and the surface in the graph show the smoothing effect when working with logarithms; which increases the region support \(S_P\).

Bootstrapping
The total number of children in the data set is 3,356,541. Each individual had a number of possible matches between zero and 3,356,540; everyone but her. This suggests a computational challenge because one will end up with 3.316 calculations. As solution, I did a bootstrap sampling method to use the empirical distribution from data, which treats the sample as being the population. The general bootstrap algorithm is as follows (Cameron and Trivedi, 2005):

1. “Given data \(b_1, b_2, \ldots, b_N\), draw a bootstrap sample of size \(N\);

2. Calculate an appropriate statistic using the bootstrap sample. Examples include (a) the estimate \(\hat{\theta}^*\) of \(\theta\), (b) the standard error \(s_{\hat{\theta}}\) of the estimate \(\hat{\theta}^*\), and (c) a t-statistic
\[ t^* = \left( \hat{\theta}^* - \hat{\theta} \right) / \hat{s}_{\hat{\theta}} \] centered at the original estimate \( \hat{\theta} \). Here \( \hat{\theta}^* \) and \( \hat{s}_{\hat{\theta}} \) are calculated in the usual way but using the new bootstrap sample rather than the original sample.

3. Repeat steps 1 and 2 \( B \) independent times, where \( B \) is a large number, obtaining \( B \) bootstrap replications of the statistic of interest, such as \( \hat{\theta}_1^*, \ldots, \hat{\theta}_B^* \).

4. Use these \( B \) bootstrap replications to obtain a bootstrapped version of the statistic\(^8\).

To estimate ITT values using the bootstrap, I took samples of \( N = 10,000 \) iid (independent identically distributed) and \( B = 3,000 \). Results are the distributions of the ITT estimated for each sample.

4. Results Section: Expected School Attendance Rate and Policy Estimators

The following section describes the results of the model of Action Arena: \textit{Más Familias en Acción}. The values of school attendance rate \( (s) \) were estimated using R as an statistical package (Zuur, Leno, 2009). The withdrawal of the Program plays with the incentive and is seen as an exogenous shock to the family income distribution \( (y) \) and children wage \( (w) \) as shown below.

\[ \text{Figure 7: Exit Strategy and Monetary Effects} \]

\[ \]

The basic empirical hypothesis is about the effects under the exit of the Program. Then, it is expected to have a negative effect by theoretical construction. Since the ITT is the average

\( ^8 \text{Cameron and Trivedi, 2005, p. 360} \)
effect; the intuition of its effects are a positive ITT effect shows probability change of family
$i, j$ to effectively send kids to school, and a negative ITT effect means a decrease in probability
by family of sending kids to school ceteris paribus. According to CTP’s theory of change, if the
school attendance rate decreases it will produce a positive effect on child labor as a compensation
reaction and vice versa. In the medium term, over more than a year, it will affect negatively
the educational achievement, and will increase the educational gaps exponentially at different
rates. In the long term, it will impact negatively the multidimensional poverty index. The
average effect ITT of school attendance probability within the observed population that lies in
$S_p$ is expected to be heterogeneous according to socio-demographic observables, therefore sub-
population analysis brings a lot of decision scenarios. This information is relevant for public
policy purposes. In sum, according to Más Familias en Acción data, the ITT means that the
probability to achieve at least 80% of school attendance rate. When $ITT > 0$ it means an
increase in the probability for a kid from a sub-population regarding the threshold (achieving
80% in the school attendance rate). When $ITT < 0$ it means a decrease in the probability of
achieving the CTP’s threshold.

The overall result of the ex ante evaluation algorithm gathers information about the 3,000
bootstraps. The resulted distribution shows a total mean value of ITT at -1.07%. This is
a small negative effect, very close to zero. So the policy withdrawal has almost no effect in
school attendance rate. The lowest percentiles of the distribution show the most dramatic
withdrawal effects (-22.27%) at the 5% of the distribution meaning that in the worst case
scenario a kid will decrease 22.27% its probability to achieve the school attendance rate’s
threshold. Those represent the most vulnerable cases, although are few observations and not
where the probability mass concentrates. This fact is seen when at the 75% percentile the
ITT is -3.33%; which means that effect changes rapidly in a positive direction and at higher
percentiles negative impact approaches to zero, as shown in the Graph below.

Furthermore, heterogeneous effects are insightful for public policy makers. Since poverty
reduction is the main social function at stake, and related to it the school attendance rate as
one of the key triggers for combating poverty in developing countries; the cost/effective criteria
force decision-makers to think on where the incentives ($r$) have the best internal returns. In
this ex ante evaluation, I see that incentives are having a very small positive effect over ($s$) in
average. However any exit strategy will have to be careful with the vacuum it will create when
it leaves; and the expected negative effects. It means, since theoretically the withdrawal of the
program will have a negative impact in the school attendance rate, the policy maker will have
to focus his attention in the most vulnerable population so can wisely minimize the negative
impact. As mentioned above, key decision dimensions at policy level are age, schoolgrade,
group, sector and gender9.

9Although in Más Familias en Acción, gender is not a policy dimension, in many other CTPs around the
Those policy dimensions provide the insights to analyze ITT by the corresponding (sub) populations. The following graphs show the heterogeneous effects on different policy relevant sub populations. The bottom line represents the 5% of the distribution (it means the lowest values of the distribution), the middle line shows the mean and the upper line represents the 95% of the distribution.

In case of the absence of subsidy, ITT mean effects are worse for girls (-1.28%) than boys (-0.87%). The mean effects regarding age show even a positive effect in early ages (5-8), at about 9 years old the impact starts to be negative, having the lowest point at the age of 13 to 15...
years old when it starts to recover again to reach almost zero effect at the age of 18. The ITT mean effects regarding school grade show similar effects with positive impact in early grades 0-2 and by third grade outcomes were almost zero decreasing constantly over the school grades reaching its highest point at 11th grade (-4.06%). It is policy relevant to see the worst-case scenario, that is for school grade is at 11th grade with (-12.33%). When analyzed by Group, it means by different incentive levels: group 1 seems to be more sensitive to withdrawal since it decreases its mean probability value to achieve the school attendance threshold by -1.95%. This policy dimension has a very disperse effects between groups and in the worst case scenario might reach even a decrease of -12.97% in its probability in Group 1. The Group 4 seems to have the least negative effects. The sector dimension confirms the rural areas are the most sensible regions with a negative average impact (-2.01%).

5. Discuss Results within an Action Arena

According to IADF there are some key context aspects to explore here. On one hand, there is going to be a natural friction with political parties at the municipal and congressional levels since they might use the social subsidies with clientelistic purposes, being reluctant to withdraw the program arguing social function goals in public but electoral interests in private. On the other hand rules-in-use might be affected, that is ex-ante evaluation might show that the subsidies have caused a permanent shock in household behavior and are not effective anymore for shaping decision making at household level, i.e. they will tend to send kids to school either with or without the incentive.

The CTP pattern of interaction described under the IADF, foresee that household effects are
going to be negative regarding school attendance rate sensible to income and child labor wage levels. Using this mechanism different effects among (sub) populations were shown at the operational level. Other set of side effects need to be taken into account. Program withdrawal will cut an established bureaucracy (set of program officers at municipal and national levels) that have a natural cost of operation. It will sink the monetary transaction costs imposed by the banking system represented in cash transfers to households. Eventually, it will represent a dereese in Minister of Education operation costs, facilities, teachers and administrative staff. At the collective choice level, a withdrawal of the Cash Transfer Program represents a challenge for the government mandate regarding accomplishing the Social Function targets. Social policy budgets might be cut-off creating a need to reallocation of resources. As said, school attendance rate decrease have a theoretical chain effect in several poverty variables, such as educational gap, child labor, illiteracy and the corresponding multidimensional poverty index. However, the ex-ante evaluation shows that a withdrawal has a minimum effect (almost none) on poverty levels, but with heterogenous effects a clear-safe reduction might be wisely displayed in order to minimize vacuum impacts of sudden withdrawal. The public policy maker’s question is about where money has better internal returns in terms of the Social Function. From the poverty standpoint a cost/benefit analysis suggests that CTP is barely having effects on poverty conditions. So invested resources in CTP are not having the desired results in terms of poverty conditions as the social function. Therefore results suggest that governmental offices might update their current poverty fighting strategies in order to maximize their desired outcomes and impacts in terms of poverty conditions. For instance, should policy makers now invest that money to improve quality standards in order to reduce poverty? i.e., better facilities and educational environments so kids have better education standards and therefore increase other poverty dimensions or variables such as housing and labor conditions. There might be alternative mechanisms where the money produces better returns in terms of poverty.

6. Conclusions

An ex ante evaluation is a public policy tool that entails several analytical techniques that together seek to (a) describe a public policy arrangement as a social choice problem; in which I suggested the use of IADF as game setter; (b) formalize the problem so a causal model can be derived with the pros (sophisticated models allow for simplifying complex problems) and cons (abstraction always takes off from reality) that come with it; (c) estimated simulated effects of hypothetical measures that allow decision makers to understand possible effects and different scenarios. Ex ante evaluations differ from ex post impact evaluations not only from its timing (before and after a decision is made) but also from the information available (variables haven’t been realized yet versus already realized ones). Key decision makers can be confident that evidence is robust and stands up for credibility, reliability and objectivity.

Ex ante evaluations are highly sensitive to formal models, they have several virtues; the ca-
sation mechanisms provide powerful explanatory capacity useful for decision makers often im-
bracht in complex situations and levels (constitutional, collective choice and operational sit-
uations); on the other hand formality brings rigour by construction (not always accuracy
and realism) which might reduce spontaneous intuition when making decisions that involve
large populations and effects at scale; and lastly low cost evaluations since require only the
current set of data from which statistical techniques can be tested to sink data gathering and
opportunity costs.

7. References

  Evolución Reciente” en Colombia Internacional 62, jul-dic 2005, 132-145

  Treatment Effects.” Econometrica 62: 467-475

- Angulo R.C., Díaz Y., Pardo R. (2011) Índice de Pobreza Multidimensional para Colom-
  bia (IPM-Colombia) 1997-2010, Archivos de Economía, Departamento Nacional de Planeación,
  Documento 832, noviembre de 2011

- Attanasio O., Meghir C., Vera-Hernández et al (2004), Baseline report on the evaluation
  of Familias en Acción, IFS Reports, Abril 2004

- Attanasio, Gómez (2012), Impactos de Largo Plazo del Programa Familias en Acción en
  Municipios de Menos de 100 Habitantes en los Aspectos Claves del Desarrollo Humano,
  IFS, Econometría, SEI.

- Attanasio O., Fitzsimons E., Gomez A. (2005), The Impact of a Conditional Cash Transfer
  Programme on Consumption in Colombia, EDePo Centre for the Evaluation of Develop-
  ment Policies, The Institute for Fiscal Studies, Report Summary Familias 01, March 2005

- Arrow, Kenneth J. (1951a) "Mathematical Models in the Social Sciences." In H.D. Las-
  well and D.T. Lerner, eds., Policy Sciences in the United States, Stanford University
  Press, 1951, pp. 129-154


  Methods Thesis, Columbia University

- Baez J., Camacho A. (2011) Assessing the Long-term Effects of Conditional Cash Trans-
  fers on Human Capital: Evidence from Colombia. S.D.
- Bernal R., Peña X., (2011) Guía Práctica para la Evaluación de Impacto, Universidad de los Andes, Facultad de Economía, Centro de Estudios sobre Desarrollo Económico (CEDE), Bogotá

- BID, 2010; Familias en Acción; Evaluación de Impacto de un Programa Piloto en Medellín, Colombia.


- Cecchini, S., Madaariaga, A. (2011) Programas de transferencias condicionadas: balance de la experiencia reciente en América Latina y el Caribe, CEPAL, NU CEPAL, SIDA


- DANE-DNP (2012) Misión para el empalme de las series de empleo, pobreza y desigualdad - MESEP, Resultados de la segunda fase, Bogotá


• Helo, J. (2008) “Una Evaluación de los Programas de Estabilización para la Población Desplazada en Colombia” Manuscrito, Universidad de los Andes, Bogotá


• Lozano, F (2012), Teoría Microeconómica: Elección Racional, Universidad Nacional de Colombia, Bogotá


• North, D. (1990), Institutions, Institutional Change and Economic Performance, Cambridge University Press

• North, D. (2005), Understanding the Process of Economic Change, Princeton University Press


Appendix A: Rationality

The utility index as a quantitative variable associated with a decision problem. It represents a numerical value for each of the consequences behind the decision problem face by an individual. The highest numbers are given to the preferred consequences (Abitbol & Botero, 2005; Arrow, 1951b; Lozano, 2012). If the alternatives have a probability, then uncertainty decision and represented by the expected utility function. The concept of rationality has been written in an axiomatic form as the rational choice model has two critical assumptions: completeness and transitivity over the preferences (Neumann, J., & Morgenstern, O. 1953). The basic model assumes perfect information and complete. She chooses an alternative after an optimization of the preferences according to alternatives and probabilities. The elements that make up the model are (Osborne & Rubinstein, 1994 p.4):

- A set of actions \((A)\) from which make decisions
- A set of possible outcomes \((C)\) over those actions
- An outcomes function \(g : A \rightarrow C\) that links consequences with actions
- An ordered preferences relationships (must be complete and transitive) over the set \(C\)

A complete and transitive relationship must meet (Lozano, 2012):
Be, \(\mathbb{R} = \{x \in \mathbb{R} \mid x \geq 0\}\)
Completeness
The relation \(\geq\) over \(\mathbb{R}\) is complete if \(\forall x^1, x^2 \in \mathbb{R}\) then \(x^1 \geq x^2\) or \(x^2 \geq x^1\) or both
Transitivity
The relation \(\geq\) over \(\mathbb{R}\) is transitive if \(\forall x^1, x^2, x^3 \in \mathbb{R}\) such that \(x^1 \geq x^2\) and \(x^2 \geq x^3\) then \(x^1 \geq x^3\)

\(^{10}\)Known as the Von Neumann-Morgenstern Utility Function
Appendix B. Kernel Density Function

Following Hansen (2009) in statistics, Kernel Density Estimation is a non-parametric way to estimate the probability density function of a random variable. Kernel density estimation is a fundamental data smoothing problem where inferences about the population are made, based on a finite data sample. Non-parametric means infinite-dimensional; its methods make the complexity of the fitted model depend upon the sample. The more information is in the sample (i.e., the larger the sample size), the greater the degree of complexity of the fitted model. Non-parametric theory acknowledges that fitted models are approximations, and therefore are inherently misspecified. Misspecification implies estimation bias. These methods typically involve some sort of approximation or smoothing method. Some of the main methods are called kernels, series, and splines.

Nonparametric methods are typically indexed by a bandwidth or tuning parameter which controls the degree of complexity. The choice of bandwidth is often critical to implementation.

Definition: Let $X$ be a random variable with continuous distribution $F(x)$ and density $f(x) = \frac{d}{dx} F(x)$. The goal is to estimate $f(x)$ from a random sample $\{X_1, \ldots, X_n\}$. The distribution function $F(x)$ is naturally estimated by the Empirical Distribution Function (EDF) $\hat{F}(x) = \frac{1}{n} \sum_{i=1}^{n} 1(X_i \leq x)$. Consider a discrete derivative. For some small $h > 0$, let

$$\hat{f}(x) = \frac{\hat{F}(x+h) - \hat{F}(x-h)}{2h}$$

we can write this as

$$\frac{1}{2nh} \sum_{i=1}^{n} 1(x + h < X_i < x + h) = \frac{1}{2nh} \sum_{i=1}^{n} 1\left(\frac{|X_i - x|}{h} \leq 1\right)$$

and the general case is

$$\hat{f}(x_0) = \frac{1}{nh} \sum_{i=1}^{n} k\left(\frac{X_i - x_0}{h}\right)$$

where $K(s)$ is a biweight kernel function

$$K(u) = \begin{cases} \frac{15}{16}(u^2 - 1)^2 & \text{if } |u| \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

The estimator $\hat{f}(x)$ counts the percentage of observations which are close to the point $x$. If many observations are near $x$, then $\hat{f}(x)$ is large. Conversely, if only a few $X_i$ are near $x$, then $\hat{f}(x)$ is small. The bandwidth $h$ controls the degree of smoothing.
The **probability density function**, the kernel function \( k(u) : \mathbb{R} \rightarrow \mathbb{R} \) is any function which satisfies that \( \hat{f}(x) \) integrates to one: 
\[
\int_{-\infty}^{\infty} \frac{1}{h} k \left( \frac{x - x_i}{h} \right) dx = \int_{-\infty}^{\infty} k(u) du = 1
\]

Change of variables: \( u = \left( \frac{x - x_i}{h} \right) \)

The **moments** of a kernel are \( k_j(k) = \int_{-\infty}^{\infty} u^j k(u) du \)

A **non-negative** kernel satisfies \( k(u) \geq 0 \) for all \( u \).

A symmetric \( k(u) = k(-u) \) for all \( u \).

Symmetric non-negative second-order kernels are **high-order kernel**

Asymptotic consistency of this estimator requires that the smoothing parameters satisfy
\[
nh_n^{\hat{w}} h_n^{\hat{w}} \rightarrow \infty; h_n^{\hat{w}} \rightarrow 0; n \rightarrow \infty
\]

The nonparametric estimator is only defined at points where the data density is positive. For this reason, we need restrict the estimation to points of evaluation that lie within the region \( S_p \) where \( S_p = \{ (\hat{y}, \hat{w}) \text{ tal que } f_{\hat{y}, \hat{w}}(\hat{y}, \hat{w}) > 0 \} \) and \( f_{\hat{y}, \hat{w}}(\hat{y}, \hat{w}) \) is the density. Todd and Wolpin (2008) determine empirically whether a particular point of evaluation \((y_0, w_0)\) lies on \( S_p \) by estimating the density at each point and checking whether it lies above a cut-off trimming level \( q_\alpha \) that is small and positive. That is, need to check whether

\[
\hat{f}(y_0, w_0) > q_\alpha,
\]

where \( \hat{f}(.) \) is a nonparametric estimate of the density. The cut-off level \( q_\alpha \) corresponds to the 2% quantile of the positive estimated density values.
Appendix C: Figures

Figure No. A: Theory of Change: Educational Cash Transfer

Adapted from Zall & Rist, 2004
Figure B: Poverty Conditions

From the IPM Colombia 1997-2008 (DNP, 2011)
Figure C: Conditional Cash Transfer Programs in the World: 1997 and 2008

1997

2008

Source: http://gpo.worldbank.org/RFYYFBQPU0
Figure D: Research Methodology

Figure E: Institutional Analysis and Development Framework

Source: Ostrom, 2005
Figure F: Economics of Institutions and Levels of Analysis

L1: Embeddedness: informal institutions, customs, traditions, norms, religion

L2: Institutional Environment: formal rules of the game – esp. Property (polity, judiciary, bureaucracy)

L3: Governance: play of the game – esp. Contract (aligning governance structures with transactions)

L4: Resource allocation and employment (prices and quantities; incentive alignment)

Meta - Constitutional Situations
Constitutional Rules-in-use
Constitutional Situations
Collective Action Rules-in-use
Collective Action Situations
Operational Rules-in-use
Operational Situations

Frequency (Years)

$10^2 - 10^3$

$10 - 10^2$

$1 - 10$

Permanente

L1: social theory
L2: Economics of property rights/positive political economy
L3: transaction cost economics
L4: neoclassical economics/agency theory

Adapted from Williamson, 2000 and Ostrom, 2005
Figure G: Action Arena: Familias en Acción

Based on Ostrom, 2005

Exogenous Variables

- Physical/Material Conditions
  - Municipalities where facilities and teachers were enough to receive school enrollment increase

- Attributes of Community
  - Families under poverty condition
  - Kids between 5 to 18 years old
  - School attendance rate is worse in rural areas

- Rules-In-Use
  - Families tend to send their kids to work to increase their income

Action Arena: Familias en Acción

Policy-making Level

Stakeholders: DPS, Minister of Education, Banking Entities.
Positions: decisions at the policy-making level, setting conditions or parameters of the game.
Outcomes: in terms of human capital and poverty indices.
Cost-Benefit ratio will be the Cost of the Program versus the Poverty Index.

Operational Level

Stakeholders: municipal liaison, teachers, public schools, households, children.
Positions: program implementation, teaching and controlling, receiving cash if accomplished, receive education and do well.
Outcomes: school rate attendance, human capital, poverty.
Cost-Benefit ratio: families get cash, kids get education, government decreases poverty and increases human capital.

Patterns of Interaction

Families’ Choice Modal for Schooling

Outcome: School Rate Attendance

Feedback Relationships

Multidimensional Poverty Index (MPI): Illiteracy - Educational Gap - Educational Achievement - Child labor