WP 2006-09
January 2006

Working Paper

Department of Applied Economics and Management
Cornell University, Ithaca, New York 14853-7801 USA

Does Conditionality Generate Heterogeneity and Regressivity in Program Impacts? The Progresa Experience

Juan Carlos Chavez-Martin del Campo

Does Conditionality Generate Heterogeneity and Regressivity in Program Impacts? The Progresa Experience

Juan Carlos Chavez-Martin del Campo*
Department of Applied Economics and Management
Cornell University
jcc73@cornell.edu

January 18, 2006

Abstract

We study both empirically and theoretically the consequences of introducing a conditional cash transfer scheme for the distribution of program impacts. Intuitively, if the conditioned-on good is normal, then better-off households tend to receive a larger positive impact. I formalize this insight by means of a simple model of child labor, applying the Nash-Bargaining approach as the solution concept. A series of tests for heterogeneity in program impacts are developed and applied to Progresa, an anti-poverty program in Mexico. It can be concluded that this program exhibits a lot of heterogeneity in treatment effects. Consistent with the model, and under the assumption of rank preservation, program impacts are distributionally regressive, although positive, within the treated population

JEL Classification: H430, C140, C210.
Keywords: *Heterogeneous Program Impacts, Regressivity, Progresa, Conditional Cash Transfers, Nonparametric Methods, Semiparametric Methods.*

*I am very thankful to Francesca Molinari and participants at the Brown-Bag Seminar at the Deparment of Applied Economics and Management, Cornell University, for their feedback and comments. My research has been supported by the Consejo de Ciencia y Tecnologia del Estado de Aguascalientes (CONCYTEA), the Consejo Nacional de Ciencia y Tecnologia (CONACYT), and the Ford/MacArthur/Hewlett Program of Graduate Fellowships in the Social Sciences. All remaining errors are my own.

1 Introduction

Nowadays, conditional cash transfer schemes (CCTS) constitute a key element of many anti-poverty programs around the world. Following Das, Do, and Oler (2004), a conditional cash transfer scheme can be broadly defined as "...any scheme requiring a specified course of action in order to receive a benefit as a conditional cash transfer". Examples of programs implementing CCTS are Oportunidades in Mexico, Red de Proteccion Social in Nicaragua, and Bolsa Familia in Brazil. The aim of these programs is to alleviate today's poverty by transferring money to poor families, and to short-circuit tomorrow's, by making the transfers conditional. The conditionality usually operates through lower bounds on human capital investment which takes the form of requiring a minimum attendance rate to school, and constant health monitoring for the children.

What is the rationale behind imposing a conditionality to the beneficiaries of a social program? if individuals are rational, there are no externalities, and policymakers have full information, then there is no case for implementing CCTS. However, these conditions are rarely met. If individuals are not fully rational, imposing a conditionality may help them to increase their own welfare. For instance, if a beneficiary is time inconsistent, then if may be optimal to impose the condition that transfers should be received in several payments.[1] If information is asymmetric in the sense that the policymaker does not have some relevant information of the beneficiaries such as income and asset holdings, then CCTS can be used as a screening mechanism with the specific purpose of improving the targeting efficiency of the program. For example, if the conditioned-on good is inferior, richer households are more likely to be screened out of the program (Besley and Coate 1992).

There is a third rational for CCTS. In the presence of externalities, individuals do not internalize the effect of their choices on others. By imposing a conditionality,

[1] See chapter 2 in this dissertation.

policymakers may be able to move individuals towards a more efficient equilibrium. One notorious case is that of child labor and human capital investment: parents usually decide children's time allocation between education and work. Since the economic benefits of child labor are immediately felt, and the economic benefits of education are only feasible in the long run, parents may not internalize the benefits of human capital investment in their children. Therefore, CCTS can be used in this case to restore efficiency by imposing lower bounds on variables such as school attendance.

Although CCTS may help policymakers to reach a more efficient economy and to reduce poverty in the long run by increasing investments on human capital today (so the children of the poor may escape poverty in the future), they could imply a tradeoff between the equity and efficiency goals of policymakers, at least in the short run. In particular, if the conditioned-on good is normal, then worse off households may be receiving less "effective" transfers than other groups of beneficiaries if participating in the program imposes some sort of opportunity cost such as foregone wages from child labor.

A good example of this tension is The Female Stipend Program in Bangladesh. This program gives stipends to girls who attend at least 85% of classes at a secondary school level with the explicit goal of increasing investment on human capital. All girls can participate in this program independently of their socioeconomic background. Since education is usually a normal good, richer households are more likely to enroll their daughters in secondary schools than households in the low tail of the income distribution. Besides, the opportunity cost of enrolling a child into school or making the 85% lower bound on school attendance is more likely to exceed the benefits obtained from the stipend for the poorest households. Khandker et al (2003) notice that the "...untargeted stipend disproportionally affects the school enrollments of girls from households with larger land wealth. Targeting towards the land poor may reduce the overall enrollment gains of the program while equalizing enrollment effects across landholding classes."

Despite the potential effects that CCTS programs have on the distribution of treatment effects, most existing research on program evaluation of anti-poverty programs focuses on mean impacts. There are, however, some studies for developed countries that take the issue of heterogeneity in program impacts into account. In an excellent study about heterogeneity in program impacts, Heckman, Smith and Clements (1997) find strong evidence of heterogeneous impacts when evaluating the US Job Training Partnership Act. In a similar spirit, Bitler, Gelbach, and Hoynes (2003) study the Connecticut's Job First program; they conclude that this welfare program exhibits a lot of heterogeneity in program impacts, just as predicted by standard labor supply theory.

In this paper, we study the distribution of program impacts in Progresa, recently renamed Oportunidades. This anti-poverty program was introduced by the Mexican government in 1997 and provides conditional cash transfers to poor families. Similar to The Female Stipend Program in Bangladesh, the conditioned-on good is school attendance which, not surprisingly, is a normal good in the case of Mexico (Lopez-Acevedo and Salinas 2000). We take advantage of the experimental design of the evaluation sample to identify the parameters of interest for this study.

Our empirical findings can be summarized in two main points. First, there is strong evidence that heterogeneity in program impacts is a common phenomenon in Progresa. Second, under the assumption of perfect positive dependence, and consistent with the model developed in this paper, better off households tend to receive larger positive program impacts than poorer households.

The paper is organized as follows. Section 2 describes Progresa, the evaluation sample, and the selection of beneficiaries. Section 3 develops a simple household bargaining model of child labor and human capital accumulation, and discusses its connection with CCTS. Section 4 briefly analyzes the evaluation problem and presents average treatment effects of the program as a benchmark case. Section 5 develops some tests for homogeneity in program impacts. Section 6 imposes a specific type of monotonicity assumption: rank preservation, and makes this assumption operational through the estimation of quantile treatment effects (QTE). Section 7 concludes. Mathematical details, algorithms, and proofs are in the Appendices.

2 Progresa

In 1997, the Mexican government introduced the Programa de Educacion, Salud y Alimentacion (the Education, Health, and Nutrition Program), better known as Progresa, and recently renamed Oportunidades, as an important element of its more general strategy to eradicate poverty in Mexico. The program is characterized by a multiplicity of objectives such as improving the educational, health, and nutritional status of poor families.

Progresa provides cash transfers, in-kind health benefits, and nutritional suplements to beneficiary families. Moreover, the delivery of the cash transfers is exclusively through the mothers, and is linked to children's enrollment and school attendance. The conditionality works as follows: in localities where Progresa operates, those households classified as poor with children enrolled in grades 3 to 9 are eligible to receive the grant every two months. The average bi-monthly payment to a beneficiary family amounts to 20 percent of the value of bi-monthly consumption expenditures prior to the beginning of the program. Moreover, these grants are estimated taking into account the opportunity cost of sending children to school, given the characteristics of the labor market, household production, and gender differences. By the end of 2002, nearly 4.24 million families (around 20 percent of all Mexican households) were incorporated into the program. These households constitute around 77 percent of those households considered to be in extreme poverty.

2.1 Data: A Quasi-Experimental Design

Because of logistical and financial constraints, the program was introduced in several phases. This sequentiality in the implementation of the Progresa was capitalized by randomly selecting 506 localities in the states of Guerrero, Hidalgo, Michoacan, Puebla, Queretaro, San Luis Potosi and Veracruz. Of the 506 localities, 320 localities were assigned to the treatment group and the rest were assigned to the control group. In total, 24,077 households were selected to participate in the evaluation sample. The first evaluation survey took place in March 1998, 2 months before the distribution of benefits started. 3 rounds of surveys took place afterwards: October/November 1998, June 1999, and November 1999. The localities that served as control group started receiving benefits by December 2000. For the empirical application of the methodologies developed in this paper, we will make use of the June 1999 round.

2.2 Progresa's selection of localities and beneficiary households

Progresa's methodology to identify potential beneficiaries consists of two two main stages: (1) the selection of localities; (2) the selection of beneficiary households within selected localities. For the first stage, a marginality index was constructed for each locality in Mexico. Based on this index, localities deemed to have a high marginality level and with more than 50 and less than 2,500 inhabitants were considered priorities for the program. Finally, budgetary constraints as well as program components that require the presence of school and clinics for the implementation of the program were considered to select the group of localities to be covered by the Progresa. For the second stage, a census, ENCASEH (Encuesta de Caracteristicas Socioeconomics de los Hogares), was conducted in each of the selected localities. Using this data, a measure of monthly per capita income per household was constructed subtracting child income from total household income. A poverty line of 320 pesos per capita per month was employed to create a new binary variable taking the value of 1 if household's monthly per capita income was below 320 pesos and 0 otherwise. Finally, discriminant analysis was employed for each geographical region. By doing so, it was possible to identify the variables that discriminate best between poor and non-poor households, and a rule to classify households as poor or non-poor was developed by estimating a discriminant score for each household.

3 A Simple Model of Human Capital Investment and CCTS

In this section, we present a simple model of child labor and human capital investment. Our objective is to shed some light on the connection between these variables and CCTS. We build this model as an extension of Baland and Robinson (2000), although we do not adopt a unitary view of the household. Similar to Kanbur and Haddad (1997) and Martinelli and Parker (2003), we adopt a bargaining perspective for the intra-household resource allocation problem.

3.1 One-Sided Altruism

We consider a one-good economy. The single good in this economy is produced with the linear technology

$$Y = L \tag{1}$$

where L is labor input measured in efficiency units of labor. We assume that the labor market is perfectly competitive.

There is a continuum of households who live for two periods, $t = 1, 2$. Each of these households is composed by a man, a woman, and a child. We will refer to the man and the woman together as the parents for the rest of the section. In period 1, parents are characterized by their income generating ability a, where a also represents efficiency units of labor. We assume that households are distributed uniformly on $[\underline{a}, \overline{a}]$, with each household inelastically supplying a efficiency units of labor per period.

In period 1, the child is endowed with one unit of time. Parents decide how to allocate the child's time between child labor, l, and human capital accumulation, h. They also decide how much to leave as a bequest to the child, b. For the sake of simplicity, it is assumed that l is measured in efficiency units of labor, so the child is endowed with one efficiency unit of labor in period 1. In the second period, the child's income generating ability is given by $\phi(h)$, where $h = 1 - l$ and $\phi(\cdot)$ is C^2, strictly increasing, and strictly concave function defined on $[0, 1]$, with $\phi(0) = 1$, $\phi'(1) < 1$, and $\phi'(0) > 1$. This technology implies that the efficient investment level on human capital, h^o, is given implicitly by $\phi'(h^o) = 1$.[2]

Let (x_{1f}, x_{2f}) and (x_{1m}, x_{2m}) denote the consumption levels of the father and the mother for periods 1 and 2, respectively. The child is assumed to consume only in period 2, with consumption level denoted by x_c. The woman cares only about her own consumption and the consumption of the child; similarly, the man cares only about his own consumption and the consumption of the child. The father's preferences are represented by

$$W_f = \alpha(\ln x_{1f} + \ln x_{2f}) + (1 - \alpha) \ln x_c \tag{2}$$

and the mother's preferences are given by

$$W_m = \beta(\ln x_{1m} + \ln x_{2m}) + (1 - \beta) \ln x_c \tag{3}$$

where $1 > \alpha > \beta > 0$.

Besides choosing the time allocation of the child, parents can also decide to make

[2] In other words, h^o is the level of human capital that maximizes the household's intertemporal income.

positive bequests to him. We denote these bequest by $b \in \mathbb{R}_+$. Parents have access to a storage technology, so they can transfer resources between periods by saving. We denote the household's saving level by s. Households are borrowing constrained in the sense that parents can save but not borrow. Therefore, parents face the budget constraints

$$x_{1f} + x_{1m} = a + l - s \tag{4}$$

$$x_{2f} + x_{2m} = a + s - b, \tag{5}$$

and

$$x_c = \phi(1 - l) + b \tag{6}$$

Decisions about $x_{1f}, x_{2f}, x_{1m}, x_{2m}, x_c, b$, and s are made by the parents in the first period by solving a *generalized Nash bargaining problem* with solution given by the following program

$$Max(x_{1f}^{\alpha} x_{2f}^{\alpha} x_c^{1-\alpha} - u_f)^{\gamma} (x_{1m}^{\beta} x_{2m}^{\beta} x_c^{1-\beta} - u_m)^{1-\gamma} \tag{7}$$

The paremeter $\gamma \in (0,1)$ introduces asymmetry into the model. The ratio $\frac{\gamma}{1-\gamma}$ can be interpreted as as the relative bargaining power of the father with respect to the mother. u_f and u_m are referred to as threat points or disagreement points. For the rest of the analysis we assume $u_f = u_m = 0$.

Proposition 1 *If savings and bequests are interior, then parents are investing the efficient level of human capital on the child. Moreover, human capital is a normal good.*

PROOF: See Appendix.

3.2 Two-Sided Altruism

We now introduce a particular form of altruism from children to parents. We will show that the results we obtained above can be extended to this new setting. We assume that children derive utility both from consumption in the second period and from any transfer to their parents:

$$W_c = \pi \ln x_c + (1-\pi) \ln \tau^c \tag{8}$$

where x_c is child's consumption when adult, τ^c is the transfer given to the parents, and $\pi \in (0,1)$.

Household choices are timed as follows. In period 1, parents choose investment on human capital and saving. Period 2 is divided in two subperiods. In the first subperiod, they choose the level of bequests. In the second subperiod, children decide how much to transfer to their parents. Therefore, they face the following budget constraint:

$$x_c + \tau^c = \phi(1-l) + b \tag{9}$$

We solve for the equilibrium by backward induction. For the second subperiod, it is easy to show that children choose the following levels of own consumption and transfers to their parents:

$$\begin{aligned} x_c &= \pi(\phi(1-l)+b) & (10) \\ \tau^c &= (1-\pi)(\phi(1-l)+b) & (11) \end{aligned}$$

Since both parents are assumed to be forward looking, parents anticipate the effect that their current decisions have both on child consumption and the transfers received from him. Therefore, the solution to the Nash bargaining problem is given by the solution to

$$Max(x_{1f}^{\alpha}x_{2f}^{\alpha}(\phi(1-l)+b)^{1-\alpha})^{\gamma}(x_{1m}^{\beta}x_{2m}^{\beta}(\phi(1-l)+b)^{1-\beta})^{1-\gamma} \tag{12}$$

subject to the constraints

$$x_{1f}+x_{1m}=a+l-s \tag{13}$$

$$x_{2f}+x_{2m}=a+s+(1-\pi)\phi(1-l)-\pi b \tag{14}$$

Proposition 2 *In the model with two-sided altruism, if savings and bequests are interior, then parents invest the optimal level of human capital on the child. Moreover, human capital is a normal good.*

PROOF: See Appendix.

3.3 Conditional Cash Transfers: Efficiency vs Equity

We now introduce a social planner whose objective is to help households to invest the optimal amount of human capital h^o on the child. The planner implements the following policy: It provides a transfer $\bar{\tau}$ to all households that invest at least the optimal level of human capital. Formally,

$$\tau = \begin{cases} \bar{\tau} & \text{if } h \geq h^o \\ 0 & \text{otherwise} \end{cases}$$

Let $V(\bar{\tau}, a, l^o)$ denote the indirect utility of a household with income generating ability a if it accepts the conditionality imposed by the policymaker. Similarly, let $V(0, a, l^*(a))$ denote its indirect utility if it does not, where $l^*(a)$ is the optimal choice of child labor for a household that does not participate in the program. Clearly, a household will accept the conditionality if $V(\bar{\tau}, a, l^o) - V(0, a, l^*(a)) > 0$. If child labor is an inferior good, or equivalently, human capital investment is a normal good, it can be shown that this difference is increasing on a, so better off households are more likely to accept the conditionality. Since the opportunity cost of participating in the program is given by the foregone income coming from child labor, $l^*(a) - l^o$, the effective transfer received by a household with income generating ability a is given by:

$$\tau^e(a) = \begin{cases} \bar{\tau} + l^o - l^*(a) & \text{if } V(\bar{\tau}, a, l^o)) > V(0, a, l^*(a)) \\ 0 & \text{otherwise} \end{cases}$$

Clearly, effective transfers $\tau^e(a)$ are non-decreasing on a for households participating in the program.[3] Therefore, within this group, better off households tend to receive a larger positive impact from the program.

More generally, we can distinguish three types of households with choices depending on their income generating ability a. The first type of household invests less than the optimal level of human capital h^o even when the CCTS is available, so it does not receive any transfer at all. The second type of household was investing less than the optimal level h^o before the scheme was available, but increases its investment level to h^o once he becomes a beneficiary of the program. Finally, the third type of household was already investing the optimal level of human capital, so it always participate in the program since it represents a pure income transfer to the household.

4 The Evaluation Problem

Although randomization helps to answer many of the questions raised by policy-makers, there are many other questions that remain unanswered, in particular those related with the distribution of program impacts across the population of beneficiaries.

To formalize the inferential problem, let each member j of population J be exposed to a mutually exclusive and exhaustive binary set of treatments $T = \{0, 1\}$, and

[3] Given the assumptions of the model, in particular the concavity of $\phi(\cdot)$, for any level of generating ability a, l^o is a lower bound for the optimal choice of child labor: i.e. $l^*(a) \geq l^o$

have a response function $y_j(t) : T \rightarrow \mathbb{R}$ mapping treatments into outcomes. The population is a probability space (J, Ω, P) and $y(\cdot) : J \rightarrow \mathbb{R} \times \mathbb{R}$ is a random variable mapping the population into their response functions. Therefore, there exist two potential states of the world for each member j of J: $(y_j(0), y_j(1))$. Lets denote program participation by the indicator variable d_j, where $d_j = 1$ indicates program participation, and $d_j = 0$ otherwise. The analyst observes d_j, but he cannot observe $y_j(0)$ and $y_j(1))$ simultaneously. More formally, he observes $y_j = d_j y_j(1) + (1 - d_j) y_j(0)$. The fact that one cannot observe both outcomes for each individual is known as the evaluation problem.

4.1 Average Treatment Effects

Following the traditional approach in the program evaluation literature, the average treatment effect on the treated (ATE) is given by

$$\tau = E[y(1) - y(0) | d = 1] \tag{15}$$

Randomization guarantees the identification of ATE since we have $P(y_0 \mid d = 1) = P(y_0 \mid d = 0)$. In fact, it turns out that ATE can be consistently estimated under the weaker assumption that d is independent of $y(0)$.[4]

Columns 2 and 3 reports estimated mean outcomes for treatment and control samples, respectively. The first two rows concern total per capital expenditure and total per capita purchase of food items. The fourth column provides average treatment effects of the program on each of these variables. These results show that the effect of Progresa on total monthly per capita expenditure was about 26 pesos (a 15% mean effect), while its ATE on total monthly per capita food purchase was about 20 pesos. These treatment effects are statistically significant at the 1% level.

[4] To see this point, decompose the difference $E[y(1)] - E[y(0)]$ as follows $E[y|d = 1] - E[y|d = 0)] = E[Y(0)|d = 1] - E[y(0)|d = 0] + \tau = \tau$.

Table 1: Per Capita Mean Outcomes and Treatment Effects

Outcome	**Treatment** (n=6946)	**Control** (m=4098)	τ	**95% CI**
Total Expenditure	202.815	176.341	26.474 (3.410)	[20.731,32.217]
Food Purchase	148.385	128.4156	19.970 (2.028)	[16.149,23.790]

From the discussion on the program evaluation problem, we know that the identification of the joint distribution $P(y_1, y_0)$ is, in general, not possible. There is a case, however, where one can identify the distribution of program impacts $P(y_1 - y_0)$. The dummy-endogenous-variable model (Heckman 1978) assumes that

$$y_j(1) = y_j(0) + \tau$$

Defining τ as the treatment effect, this assumption implies homogeneous treatment responses. Therefore, the distribution of program impacts is the Dirac measure at τ:

$$P(y(1) - y(0) \mid d = 1) = P(\tau \mid d = 1) \tag{16}$$

Under random treatment selection, we have $\tau = E[y(1) - y(0)|d = 1]$, which is identified. Hence, the dummy-endogenous variable model identifies the distribution of program impacts.

4.2 Fréchet Space

Because of the evaluation problem, one cannot observe an individual's outcome in both treatment and control states. Therefore, it is not possible to identify the distribution of program impacts without imposing more structure on the problem at hand. However, we may be able to partially identify some features of the distribution of the random vector $(y(0), y(1))$ when $P(y(1))$ and $P(y(0))$ are identified.[5]

Let us introduce the following notation. H denotes the bi-dimensional cumulative distribution function of the random vector $(y(0), y(1))$, where $H(\mathbf{t}) = P(y_0 \leq t_1, y_1 \leq t_2)$, with $\mathbf{t} = (t_1, t_2) \in \mathbb{R}^2$. $\mathcal{H}$ denotes the *Fréchet Space* given the marginals, that is $\mathcal{H}(F_0, F_1)$ is the space of all cumulative distribution functions $H(\mathbf{t})$ on R^2 with fixed marginal cumulative distribution functions $F_0(t_1) = P(y_0 \leq t_1)$ and $F_1(t_2) = P(y_1 \leq t_2)$. We denote by E_H the expectation operator under the joint distribution H.

Frechett (1951) showed that the distribution $H(x_1, x_2)$ belongs to $\mathcal{H}(F_0, F_1)$ if and only if

$$H_-(t_1, t_2) \leq H(t_1, t_2) \leq H_+(t_1, t_2) \tag{17}$$

for all $(t_1, t_2) \in \mathbb{R}^2$, where

$$\begin{aligned} H_-(t_1, t_2) &= \max\{F_1(t_2) + F_0(t_1) - 1, 0\} \qquad (18) \\ H_+(t_1, t_2) &= \min\{F_0(t_1), F_1(t_2)\} \qquad (19) \end{aligned}$$

More recently, Ruschendorf (1981) showed that these bounds are sharp.

Tchen (1980) has established a result that will be proved to be very useful for the purposes of the present analysis. This result states that nonnegative and convex functions are monotone on the *Fréchet Space*:

Lemma 1 *(Tchen 1980) For any convex nonnegative function ψ defined on $\mathbb{R}$,*

$$E_H\psi(y_1 - y_0) \in \left[E_{H_+}\psi(y_1 - y_0), E_{H_-}\psi(y_1 - y_0)\right]$$

for all $H \in \mathcal{H}(F_1, F_2)$.

[5] For a review of the partial identification approach see Manski (2003).

4.3 Partial Identification of Mobility Treatment Effects

Because of the evaluation problem, many distributional scenarios are consistent with the data at hand. Could it be possible to "measure" this multiplicity of scenarios through some statistic? In this section we provide a way to do it by applying the same kind of logic we can find in studies of economic mobility.

While the goal of analyzing treatment effects is to predict the outcomes that would occur if different treatment rules were applied to the population (Manski 2003), the study of economic mobility centers on quantifying the movement of the units of analysis through the distribution of economic well-being over time. More precisely, research on economic mobility tries to connect past and present, "establishing how dependent one's current economic position is on one's past position..." (Fields 2001). In this sense, the analysis of economic mobility does not have to face the evaluation problem since both states, past and present, are observed in principle.

Suppose for a moment that we were able to identify counterfactual outcomes for two individuals. One of the individuals experiences a program impact of +100, the other individual experiences a decrease in the outcome of interest of -100. Keeping everything else constant, how much outcome *movement* has taken place? The standard approach to answer this question is to estimate the average treatment effect, so the net effect of the treatment is zero. After this simple exercise, one is left with the feeling that overall the treatment effect has been totally neutral. However, the fact that the two individuals considered in this simple example registered changes in their outcomes implies that the treatment is not neutral at all.

Fields and Ok (1996) define a measure of mobility that considers symmetric income movements[6] as $\int \mid w - z \mid dH(w, z)$, where w_i and z_i are the incomes of individual i at two different points on time. We can extend this measure to the context of program

[6]Symmetric outcome movement arises when individuals' outcomes change from one state to another and one is concerned about the *magnitude* of these fluctuations but not their direction.

evaluation by redefining these variables, so our measure of mobility treatment effects would be given by

$$m = \int \mid y_1 - y_0 \mid dH(y_1, y_0) \tag{20}$$

In contrast to mobility analysis, when analyzing treatment effects one has no information on counterfactual outcomes for the treated population, so we cannot identify this measure. However, we can partially identify m since the absolute value function is convex and positive (Lemma 3.1).

One complication arises since most data sets, and the Progresa data set is not the exception, have unbalanced sample sizes, that is to say, the number of observations in the treatment group is not the same than the number of observations in the control group. We circumvent this problem by using quantiles of the empirical distributions $\hat{F}_0$ and $\hat{F}_1$[7]. Table presents some estimations of m based on 100, 500, and 900 quantiles. The bootstrap confidence intervals were estimated using 2000 bootstrap replications. The ratio m_{H_+}/m_{H_-} is around six to one, which indicates that a great number of distributional scenarios are compatible with the data at hand. Because

Table 2: Mobility Treatment Effects

Number of Quantiles	$[m_{H_+}, m_{H_-}]$	95% Normal CI	90% Percentile CI
100	[25.719, 147.524]	[21.167,151.148]	[22.280, 152.076]
500	[26.461, 156.838]	[22.156,161.246]	[23.030, 160.954]
900	[26.004, 158.575]	[21.609,163.339]	[23.178, 162.985]

of the evaluation problem, one cannot discard the possibility of having an important subset of the treated population receiving negative treatment effects when ATE are strictly positive. Let $\mathcal{L} = \{j \in J : y_j(1) < y_j(0)\}$ denote the set of members of population J that register a loss as a result of participating in the program. Although it is not possible to identify the set of individuals who belong to this set in general, we

[7] The applied algorithm is described in more detail in Appendix B.

can partially identify a parameter that may shed some light on the potential negative effects of being exposed to the treatment, at least in average sense.[8] We define the average loss of participating in the program as follows[9]

$$L_H = \int \mathbf{1}_{\mathcal{L}}(y(1) - y(0))dH \quad (21)$$

$$= \int min(y(1) - y(0), 0)dH \quad (22)$$

Lemma 2 *Sharp bounds on L are given by* $[L_{H_-}, L_{H_+}]$.

PROOF: See Appendix.

Table 3: Average Loss

Number of Quantiles	$[L_{H_-}, L_{H_+}]$	95% Normal CI	90% Percentile CI
100	[-61.589,0.000]	[-64.752,0.000]	[-64.326,0.000]
500	[-66.062, -.724]	[-69.784,0.000]	[-69.083,-.004]
900	[-66.607, -.317]	[-70.356,0.000]	[-69.947,-.020]

Table 3 presents some estimations of L based on 100, 500, and 900 quantiles (the bootstrap confidence intervals were also estimated using 2000 bootstrap replications). Even though these worst case bounds may seem exaggerated at a first sight, especially the lower bound, they are a reminder that ATE may be missing a lot of relevant information. This empirically corroborates the fact that the evaluation problem generally implies the existence of multiple distributional scenarios consistent with the data generating process.

[8] Notice that these are worst case bounds. Monotonicity assumption motivated my program design and economic theory can be proved to be very helpful to improve inference.

[9] $\mathbf{1}_{\mathcal{L}}$ is the indicator function which is equal to one if $j \in \mathcal{L}$, and 0 otherwise.

5 Testing for Homogeneity in Program Impacts

In this section, we apply a partial identification approach that will allow us to develop simple tests to evaluate the hypothesis of homogeneous treatment effects on the treated.

Consider testing

$$H_0 : y(1) - y(0) = c \text{ a.s.}$$

versus

$$H_1 : y(1) - y(0) \neq c \text{ a.s.}$$

For some real number $c = E(y_1) - E(y_0)$.

Define the functional

$$\Phi(F_0, F_1) = \int \psi(y(1) - y(0))dH_+ - \psi(\int y(1)dF_1 - \int y(0)dF_0) \tag{23}$$

where $\psi(\cdot) : \mathbb{R} \to \mathbb{R}_+$ belongs to the class of nonnegative and strictly convex real valued functions. The following result will be proved to be very helpful for testing the hypothesis of homogeneous program impacts:

Proposition 3 *Let $\psi(\cdot) : \mathbb{R} \to \mathbb{R}_+$ be any nonnegative and strictly convex real valued function. If $\Phi(F_0, F_1) > 0$, then $y(1) - y(0) \neq c$ a.s.*

PROOF: See Appendix.

Therefore, we could test the hypothesis of homogeneity in program impacts through testing the hypothesis $H_0 : \Phi(F_0, F_1) = 0$. As an example, let $\mathcal{W}_\alpha$ denote the family of functionals defined by

$$\left\{\Phi_\alpha(F_0, F_1) : \Phi_\alpha = \int (|y(1) - y(0)|^\alpha dH_+ - |\int y(1)dF_1 - \int y(0)dF_0|^\alpha, \alpha \geq 2\right\}$$

It can be shown that $\psi(x) =\mid x \mid^{\alpha}$ is a strictly convex function[10] for $\alpha \geq 2$ (See Appendix). Therefore, $\Phi_{\alpha}(F_0, F_1) > 0$ implies $y(1) - y(0) \neq c$ a.s.

From here, we can derive an indirect way of testing the null hypothesis by statistically comparing the hypothesis

$$H_0 : \Phi_{\alpha}(F_0, F_1) = 0$$

versus

$$H_1 : \Phi_{\alpha}(F_0, F_1) \neq 0$$

Corollary 1 *The hypothesis of homogeneous treatment effects can be rejected if*

$$Var(Y(1)) \neq Var(Y(0))$$

PROOF: See Appendix.

Corollary 3.1 can be proved to be very helpful if we impose more structure on the problem. Let $Y_i(1) \sim N(\mu_1, \sigma_1^2)$ and $Y_j(0) \sim N(\mu_0, \sigma_0^2)$, $i = 1, \ldots, n$, $j = 1, \ldots, m$, be two independent random samples. Notice that

$$\frac{S_1^2/\sigma_1^2}{S_0^2/\sigma_0^2} \sim F_{n-1,m-1}$$

where S_i^2, $i = 0, 1$, is the sample variance, and $F_{n-1,m-1}$ is the F distribution with $n - 1$ and $m - 1$ degrees of freedom. Therefore, if the populations are normally distributed, we can test H_0 by statistically testing the hypothesis $\frac{\sigma_1}{\sigma_0} = 1$.

Table 4: F test for $H_0 : \frac{\sigma_1}{\sigma_0} = 1$

S_0^2	S_1^2	$f = \frac{S_1^2}{S_0^2}$	n	m	$P(F_{n-1,m-1} < f)$
18791.13	27779.93	1.478	6946	4098	1

From table 3.4, it is clear that we can reject the hypothesis that both populations share the same standard deviation. Consequently, the hypothesis of homogeneity

[10]Notice that $Var_{H_+}(Y(0) - Y(1))$ is a member of $\mathcal{W}_{\alpha}$ since $Var_{H_+}(Y(0) - Y(1)) = \Phi_2$.

in program impacts is also rejected. However, this test is not accurate unless the distributions of the populations are close to normal.[11]

We also apply other tests for equality of variances that are less sensitive to departures from normality. Levene's test (1960) tends to be more robust than the F test when the distribution is not Gaussian. Brown and Forsythe (1974) extended Levene's test to use either the median or the trimmed mean instead of the mean. Let W_0, $W_{.50}$, and $W_{.10}$ denote, respectively, the original Levene's statistic, the Levene's statistic replacing the mean by the median, and the Levene's statistic with a 10% trimmed mean.[12] All of these tests reject the null hypothesis at the 1% level (see table 3.5).[13]

Table 5: Levene's statistics

W_0	$W_{.10}$	$W_{.50}$
9.003	7.796	8.019

Another alternative is to use bootstrap methods to estimate the distribution of the statistic $\hat{\Phi}$ by resampling the evaluation samples (Efron 1979). Let $P(\hat{\Phi} - \Phi \mid F_0, F_1)$ denote the exact, finite sample distribution of $\hat{\Phi} - \Phi$. Using standard notation from the bootstrap literature, let $\hat{\Phi}^* - \hat{\Phi}$ be computed from observations obtained according to the empirical distributions $\hat{F}_0$ and $\hat{F}_1$ in the same way $\hat{\Phi} - \Phi$ is computed from the true observations $Y_i(1) \sim F_1$ and $Y_j(0) \sim F_0$, $i = 1, \ldots, n$, $j = 1, \ldots, m$. Finally, let Q^*_β denote the β-quantile of the CDF of $\hat{\Phi}^* - \hat{\Phi}$. That is

$$Q^*_\beta = \inf\{\hat{\Phi}^* : P(\hat{\Phi}^* - \hat{\Phi} \mid \hat{F}_0, \hat{F}_1) \geq \beta\} \tag{24}$$

[11] A skewness and kurtosis test rejects the hypothesis of normality. In fact, the hypothesis of symmetry can also be easily rejected.

[12] Brown and Forsythe(1974) reached the conclusion that using the trimmed mean performed best when the underlying data followed a Cauchy distribution (i.e., heavy-tailed) and the median performed best when the underlying data followed a (i.e., skewed) distribution. Using the mean provided the best power for symmetric, moderate-tailed, distributions

[13] The Levene's test rejects the null hypothesis if $W > F_{1,n+m-2}$ where $F_{1,n+m-2}$ is the upper critical value of the F distribution for some predetermined significance level.

A commonly applied method to test H_0 is to assume that $\hat{\Phi}-\Phi$ is normally distributed, and then to use the bootstrap estimate of the standard deviation as an approximate estimator of the true sample variance. That is

$$\frac{\hat{\Phi}-\Phi}{\hat{\sigma}^*} \sim N(0,1) \tag{25}$$

However, under the null, Φ is at the boundary of the parameter space since $\Phi \geq 0$. This implies that the random quantity $\hat{\Phi}/\hat{\sigma}^*$ is always positive, and hence it cannot be normally distributed with mean zero and variance one.

One possible solution for this problem is to follow Efron (1987) by assuming the existence of a monotone increasing transformation $\varphi(\cdot): \mathbb{R} \to \mathbb{R}$ such that

$$\varphi(\hat{\Phi}) - \varphi(\Phi) \sim N(z_\Phi, \sigma_\Phi^2) \tag{26}$$

for every choice of Φ (z_Φ is know as the bias correction term). For the purpose of the present study, we can weaken this assumption by requiring just symmetry for the distribution of $\varphi(\hat{\Phi}) - \varphi(\Phi)$.

Proposition 4 *Suppose there exists a strictly increasing function $\varphi(\cdot): \mathbb{R} \to \mathbb{R}$ such that*

$$\varphi(\hat{\Phi}) - \varphi(\Phi) \mid F_0, F_1 \sim V$$

$$\varphi(\hat{\Phi}^*) - \varphi(\hat{\Phi}) \mid \hat{F}_0, \hat{F}_1 \sim V$$

where V is continuously and symmetrically distributed about $\gamma \in \mathbb{R}$, satisfying

$$F_V(2\gamma + F^{-1}(\beta)) > 0$$

for some $\beta \in (0, 1/2)$. Then

$$\textit{Reject } H_0 \textit{ if } \min \hat{\Phi}^* > 0$$

is a level β test.

PROOF: See Appendix.

We estimate the bootstrap cdf of $\hat{\Phi}_2^*$ using $B = 2000$ bootstrap replications. From table 3.6, it can be inferred that the null hypothesis of homogeneous treatment effects can be easily rejected under the assumptions of the proposition. For instance, if $V \sim N(-\sigma\gamma, \sigma^2)$, we have

$$\begin{aligned} P(\varphi(\hat{\Psi}^*) - \varphi(\hat{\Psi})) &= P(\sigma Z - \sigma\gamma < 0) \\ &= P(Z < \gamma) \\ &= F_Z(\gamma) \end{aligned}$$

A plug in estimator for γ is therefore given by

$$\begin{aligned} \hat{\gamma} &= F_Z^{-1}\left(\frac{\#\{\varphi(\hat{\Phi}^*) < \varphi(\hat{\Phi})\}}{B}\right) \\ &= F_Z^{-1}\left(\frac{\#\{\hat{\Phi}^* < \hat{\Phi}\}}{B}\right) \end{aligned}$$

where the last equality follows from the monotonicity of $\varphi(\cdot)$. As expected, the sign of this parameter is strictly negative, taking values in the range $(F_Z^{-1}(.25), F_Z^{-1}(.50))$ for $\#q \in \{100, 300, 600, 800, 1000\}$, where $\#q$ indicates the number of quantiles used in the estimation. Therefore, we can reject the null hypothesis at the 1% level under the assumption of normality.

Andrews (2000) argues that the bootstrap may not be consistent when the parameter of interest is on a boundary of the parameter space. One possible solution is to draw subsamples of size $k < \min(n, m)$ from the original data with replacement. This sampling method is identical to the standard bootstrap in every aspect, but the

Table 6: Summary statistics for the bootstrap distribution of $\hat{\Phi}_2^*$ using empirical quantiles of F_0 and F_1.

Number of Quantiles	$Q^*_{.01}$	$Q^*_{.05}$	$Q^*_{.10}$	$Q^*_{.50}$	Mean	sd	Min$\hat{\Phi}_2^*$	$\hat{\Phi}_2$
100	36.481	54.2097	67.888	150.27	178.528	113.198	15.455	88.487
300	53.022	80.851	101.669	285.202	474.225	519.634	30.183	95.315
600	78.661	122.863	162.053	502.72	678.995	592.624	39.900	405.422
800	96.059	150.736	215.238	580.472	809.775	949.181	57.945	309.866
1000	109.783	190.4319	258.657	681.292	1060	1523.616	50.742	346.014

size of each replication.[14] Another possible advantage of this method is that we can estimate the bootstrap distribution of $\hat{\Phi}^*$ without using the quantiles of the empirical distributions as the original data. Table 3.7 presents several quantiles of the bootstrap distribution of $\hat{\Phi}_2^*$ for $k \in \{1000, 2000, 3000, 3500, 4000\}$. Under the assumption of normality and bias correction, the hypothesis of homogeneity in program impacts can be rejected at the 1% level.

6 Identification of Program Impacts under Monotonicity Assumptions

The bounds implied by the Fréchet Space of bivariate distributions proved to be very helpful for developing a test for homogeneity of program impacts. However, without further assumptions, it is an impossible task to pin down the actual distribution of treatment effects even in the case of a random experiment.

Inference on the distribution of program impacts may be improved by imposing assumptions implied by economic theory or any other mechanism related with the data generating process such as program design. Manski (1997) investigates what may be learned about treatment response under the assumptions of monotone, semi-monotone, and concave-monotone response functions. He shows that these assumptions have identifying power, particularly when compared to a situation where no prior information exists (worst case bounds). Typically, the type of monotonicity assumptions applied by econometricians dealing with partially identified parameters take some form of stochastic dominance. For instance, in a missing treatments environment, Molinari (2005) shows that one can extract information from the observations

[14]Bickel et al (1997) discuss a number of resampling schemes under which the size of the sample replication is smaller than the original sample size. They argue than the k out of n sampling scheme works very well in all known realistic examples of bootstrap failure.

Table 7: Summary statistics for the bootstrap distribution of $\hat{\Phi}_2^*$ using a $k/\min(n,m)$ resampling scheme.

k	$Q^*_{.01}$	$Q^*_{.05}$	$Q^*_{.10}$	$Q^*_{.50}$	Mean	sd	Min$\hat{\Phi}_2^*$
1000	175.041	371.705	558.809	2206.117	5502.247	6903.977	84.790
2000	239.864	411.428	556.666	3621.683	4463.259	4163.627	105.5011
3000	282.169	476.443	650.213	3327.524	4048.062	3188.16	131.8912
3500	299.809	482.117	680.626	3265.586	3951.013	2970.799	152.982
4000	302.032	486.642	754.739	3254.25	3835.516	2728.157	102.811

for which treatment data are missing using monotonicity assumptions. Specifically, one could assume that the effect of a social program on the outcome of interest cannot be negative. This is equivalent to assume that for each j in J we have

$$\tau_j = \max\{y_j(1) - y_j(0), 0\} \tag{27}$$

Given the design of Progresa, this seems to be a reasonable assumption. One can expect a positive effect on the outcome of interest, in our case consumption, for treated individuals. Moreover, this type of assumption implies first order stochastic dominance (FSD) of distribution $P(y(1))$ over distribution $P(y(0))$: i.e. $P(y_1 \leq x) \leq P(y_0 \leq x)$ for all $x \in \mathbb{R}$. Actually, this assumption is stronger than FSD. It implies that, for all $t \in \mathbb{R}$, we have

$$P(y(1) \geq t \mid y(0) = t) = 1 \tag{28}$$

Notice that the converse is not always true, that is to say, stochastic dominance does not necessarily imply monotonicity.[15]

From section 3.3, we know that if human capital is a normal good, a policy-maker implementing a CCTS faces a dilemma: on the one hand, it may represent a very helpful policy tool for achieving an efficient level of human capital. On the other hand, this policy instrument may be at odds with a more equal distribution of *effective* benefits. In particular, it was argued that when the conditioned-on good is normal, better-off households tend to receive larger "effective" benefits, once the opportunity cost of foregone earnings from child labor is deducted. Unfortunately, because of the evaluation problem, we cannot test this hypothesis without imposing more assumptions. We circumvent this problem by establishing a different type of

[15] To see that, just observe the random vector $(y(0), y(1))$ whose support consists of two points: (1,3) and (2,1). Clearly $y(1)$ stochastically dominates $y(0)$, but the monotonicity assumption is violated.

monotonicity assumption, one that will allow us to test the hypothesis of regressivity in program impacts.

We assume the existence of a non-decreasing real valued function $\phi(\cdot) : \mathbb{R} \to \mathbb{R}$ such that

$$y_j(1) = \phi(y_j(0)) \tag{29}$$

Notice that function $\phi(\cdot)$ is not indexed, and in consequence this assumption implies *rank preservation* among the members of population J. More precisely, in the contest of program evaluation, rank preservation means that, for some outcome of interest Y, the rank of a particular unit of observation i with respect to any other observation j is the same in both treatment and control states. More formally, rank preservation implies that, for any two members i and j of population J, the following relation holds:

$$(y_i(0), y_i(1)) \geq (y_j(0), y_j(1))$$

This untestable assumption is also a necessary condition for the existence of regressive program impacts[16].

Let $\tau(y(0)) = \phi(y(0)) - y(0)$. We say that there is regressivity in program impacts whenever $\tau(y(0))$ is a non-decreasing and non-trivial function of $y(0)$, that is, for any $i, j \in J$, such that $y_i(0) > y_j(0)$, we have[17]

$$\frac{\phi(y_i(0)) - \phi(y_j(0))}{y_i(0) - y_j(0)} \geq 1 \tag{30}$$

In order to make this result operational, and to test the hypothesis of regressivity in program impacts, we will use quantile treatment effects[18](QTE), which are a natural

[16] Let $(y_i(0), y_i(1))$ and $(y_i(0), y_i(1))$ be the outcomes in both states for $i, j \in J$. Without loss of generality, let $y_i(0) > y_j(0)$. Regressivity in program impacts is equivalent to $y_i(1) - y_i(0) > y_i(1) - y_i(0)$, which implies $y_i(1) > y_j(1)$

[17] More precisely, there is regressivity in program impacts if $\tau(y(0))$ is a non-decreasing and non-trivial function almost everywhere.

[18] See Koenker and Bassett (1978) for an application of quantile estimation to a regression setting.

extension of rank preservation to the analysis of distribution of treatment effects. Let us introduce this concept more formally. The q^{th} quantile of distribution $F_i(y)$, $i = 0, 1$ is defined as:

$$y_i(q) = \inf\{y : F_i(y) \geq q\}$$

The following result will be proved to be useful for the empirical application. It shows the existence of the function $\phi(\cdot)$ under some mild continuity assumption:

Lemma 3 *If $y_0(q)$ is a continuity point of F_0, then there exists a non-decreasing function $\phi(q) : (0, 1) \to \mathbb{R}$ such that $y_1(q) = \phi(y_0(q))$; moreover, there exists a unique function $\tau(q)$ satisfying $\tau(q) = y_1(q) - y_0(q)$.*

PROOF: See Appendix.

The QTE for quantile q can be defined as the difference in treatment status between quantile q of treatment group and quantile q of control group. Formally, QTE for quantile q is given by

$$\tau(q) = y_1(q) - y_0(q)$$

Therefore, QTE represent an alternative way for testing for regressivity in program impacts. A non decreasing and non-trivial QTE function is strong evidence for regressive program impacts under the assumption of rank preservation, where for rank preservation we mean rank preservation in terms of quantiles of the distributions F_1 and F_0.

Tables 3.8 and 3.9 introduce the QTE estimator for per capita total expenditures and per capita food purchase, respectively, for several quantiles. These quantiles were estimated simultaneously, so statistical comparisons can be made among them. Empirical variance of QTE was calculated by means of 200 bootstrap replications of the quantile treatment effect.

Table 8: Quantile Treatment Effects: Total Expenditure

q	.05	.10	.15	.20	.25	.30	.35	.40	.45	.50
$\tau(q)$	9.487	13.800	17.194	18.304	20.250	22.608	24.141	23.824	24.590	24.815
	(1.870)	(1.515)	(0.985)	(1.355)	(1.693)	(1.602)	(1.636)	(1.765)	(1.597)	(2.141)
q	.55	.60	.65	.70	.75	.80	.85	.90	.95	
$\tau(q)$	25.708	25.722	27.984	27.821	28.564	28.687	33.174	40.471	47.476	
	(2.120)	(2.337)	(2.499)	(2.813)	(3.077)	(3.411)	(4.533)	(6.816)	(11.740)	

Table 9: Quantile Treatment Effects: Food Purchase

q	.05	.10	.15	.20	.25	.30	.35	.40	.45	.50
$\tau(q)$	4.728	9.630	11.134	11.825	13.698	15.065	15.832	16.935	18.117	19.129
	(1.430)	(1.187)	(1.014)	(1.228)	(1.265)	(1.055))	(1.118)	(1.215)	(1.379)	(1.203)

q	.55	.60	.65	.70	.75	.80	.85	.90	.95
$\tau(q)$	20.029	20.966	22.823	22.946	26.071	27.286	28.504	33.757	39.332
	(1.397)	(1.273)	(1.637)	(2.012)	(2.117)	(2.266)	(2.678)	(3.581)	(6.754)

We plot these QTE in figures 3.1 and 3.2. For comparison purposes, we plot the average treatment effect as a horizontal dashed line. Dotted lines surrounding the ATE line represent a 95% confidence interval. Clearly, the variation of the treatment effects across the different quantiles is both economically and statistically significant, particularly at the extremes of the QTE plot, although for a broad band treatment effects are statistically homogeneous.

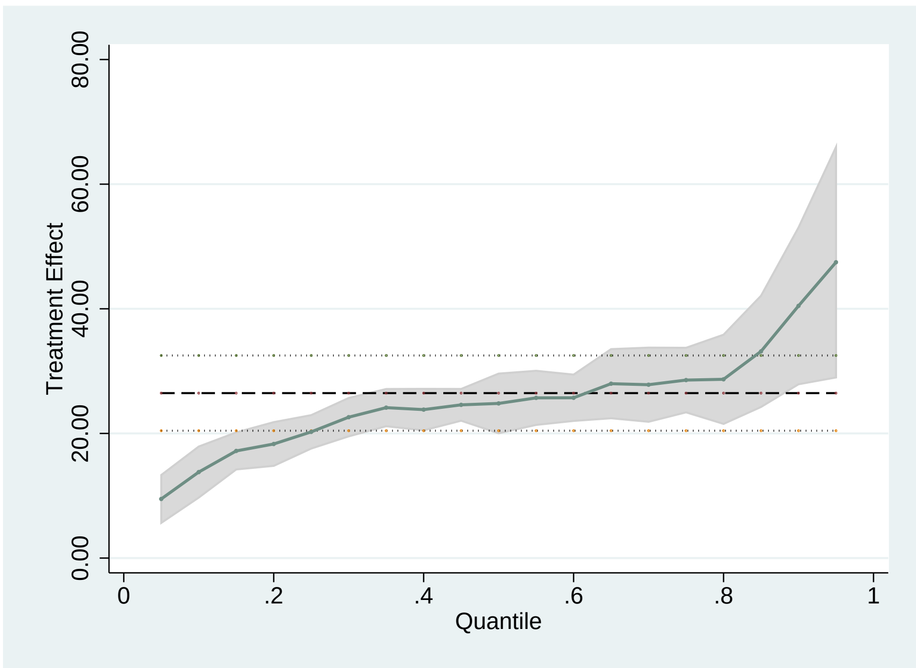

Figure 1: Quantile Treatment Effects: Total Expenditure

The QTE estimators are consistent with the predictions of the theoretical model: better off households tend to receive larger positive impacts from the program which generates the monotonically increasing shape of the QTE. This characteristic of the QTE for PROGRESA is more remarkable when one contrast que treatment effect between the lower 20 and the upper 20 centiles. For instance, in the case of total expenditure, the treatment effect for the 95th centile is about five times the treatment effect estimated for the 5th centile. This gap is about 10 times between the same

Abadie, Angrist, and Imbens (2002) extend their idea to the estimation of quantile treatment effects. See Appendix C for a description of the QTE estimator.

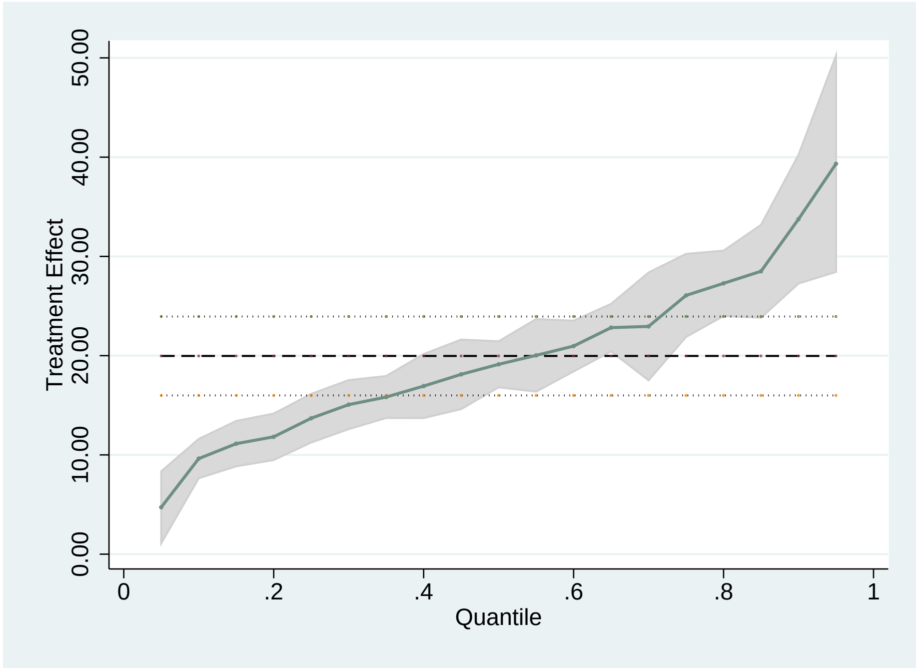

Figure 2: Quantile Treatment Effects: Food Purchase

quantiles in the case of food purchase.

QTE estimation also represents an alternative method to test for homogeneity in program impacts under some mild regularity conditions. This assertion is formalized in the following Lemma:

Lemma 4 *If treatment effects are homogeneous across the population, then* $\tau(q) = \tau$ *for all* $q \in [0, 1]$ *such that* $y_0(q)$ *is a continuity point of* F_0.

PROOF: See Appendix.

From tables 3.7 and 3.8, we can conclude that the hypothesis of homogeneity in program impacts can be rejected under the conditions of Lemma 3.4.

7 Conclusions

Conditional cash transfers represent an important policy tool for fighting poverty, particularly when there is some type of externality that prevents the poor from reaching more efficient equilibria. Human capital investment is just one example of an activity generating positive externalities. Correcting for these externalities is then an important step to break the circle of intergenerational poverty.

There are some issues, however, that should be considered by policymakers implementing this type of programs. If the conditioned-on good is normal, then it is very likely that the distributional effects of the program will be far from being distributionally neutral. In fact, as we saw in the empirical analysis, heterogeneous treatment effects are pervasive, at least in the case of the Progresa evaluation sample.

Under the assumption of rank preservation, program impacts tend to be distributionally regressive for the population participating in Progresa. As it was argued in the text, this finding is also consistent with the fact that the conditioned-on good is normal. Therefore, if the assumption of rank preservation is correct, the poorest of the poor may not be receiving as much benefits as policymakers believe they are. This has important implications for the design of antipoverty policies: policymakers should consider the existent tradeoff between equity and efficiency of outcomes in order to better understand the consequences and limitations of CCTS like Progresa. The final answer will much depend on the benefits and costs of improving the targeting efficiency of a program.

Appendix A: Proofs and Derivations

Proof of Proposition 3.1: The household bargaining model is solved through the program

$$\begin{aligned}
\mathcal{L} &= (x_{1f}^{\alpha}x_{2f}^{\alpha}(\phi(1-l)+b)^{1-\alpha})^{\gamma}(x_{1m}^{\beta}x_{2m}^{\beta}(h(1-l)+b)^{1-\beta})^{1-\gamma} + \\
&\quad \lambda_1(a+l-s-x_{1f}-x_{1m}) + \\
&\quad \lambda_2(a+s+(1-\pi)\phi(1-l)-\pi b-x_{2f}-x_{2m}) + \\
&\quad \lambda_3 s+\lambda_4 b
\end{aligned}$$

from where we can obtain the following first order conditions:

$$\begin{aligned}
\frac{z_1}{x_{1f}} &= \lambda_1 \\
\frac{z_1}{x_{2f}} &= \lambda_2 \\
\frac{z_3}{x_{1m}} &= \lambda_1 \\
\frac{z_4}{x_{2m}} &= \lambda_2 \\
\frac{-z_2\phi'(1-l)}{\phi(1-l)+b}+\lambda_1-\lambda_2(1-\pi)\phi'(1-l) &= 0 \\
\frac{z_2}{\phi(1-l)+b}-\lambda_2\pi+\lambda_4 &= 0 \\
-\lambda_1+\lambda_2+\lambda_3 &= 0
\end{aligned}$$

where $z_1 = \alpha\gamma$, $z_2 = (1-\alpha)\gamma+(1-\beta)(1-\gamma)$, and $z_3 = \beta(1-\gamma)$. From the first order conditions we have

$$\begin{aligned}
x_{1m} &= \frac{z_3}{z_1}x_{1f} \\
x_{2m} &= \frac{z_3}{z_1}x_{2f} \\
\frac{z_2\phi'(h)}{\phi(h)+b} &= \frac{z_1}{x_{1f}}
\end{aligned}$$

and

$$\phi'(h) \geq 1$$

with the last condition holding with equality if $(b, s) \in R^2_{++}$.

For the second part, assume the household is both savings and bequest constrained, so $b = s = 0$. Assume also that it receives an exogenous transfer of income $\omega > 0$ in period 1. From the first order conditions

$$\frac{\phi'(1-l)}{\phi(1-l)} = \frac{z_1}{z_2 x_{1f}} = \frac{z_3}{z_2 x_{1m}}$$

Since the household is both bequest an savings constrained, we have $l^* < l^o$ and $x_{1f} + x_{1m} = a + \omega + l$. Assume, towards a contradiction, that child labor does not decrease: i.e. $\Delta l \geq 0$. Hence, either x_{1f} or x_{1m} increases. This fact and the condition above together imply an increase in child labor, a contradiction. Therefore, human capital is a normal good. □

Proof of Proposition 3.2: The proof for the first part of the proposition is along the lines of the case with one-sided altruism. To prove that human capital is a normal good, assume the household receives an exogenous positive transfer in period 1, say $\Delta\omega > 0$. If the household is both saving and bequest constrained, then the budget constraint in period 1 is given by $x_{1m} + x_{1f} = a + \Delta\omega + l$. Assume, towards a contradiction, that child labor increases. From the first order conditions, both x_{1m} and x_{1m} increase in equilibrium. From the first order conditions, we also have:

$$\frac{z_3}{x_{1f}} = \left[z_2 + \frac{\left[z_4 \left(\frac{z_4 + z_1}{z_1} \right) \right] (1-\pi)}{a + (1-\pi)} \right] \frac{\phi'(1-l)}{\phi(1-l)}$$

This is clearly a contradiction since the left-hand side of the equation strictly decreases, while the right-hand side increases or remains constant.□

Lemma 5 *Sharp bounds on the correlation coefficient* ρ_{Y_0,Y_1} *are given by*

$$[\rho_{H_+}^{Y_0,Y_1}, \rho_{H_-}^{Y_0,Y_1}]$$

PROOF:

$$\begin{aligned}
\rho_{Y_0,Y_1} &= \frac{Var(Y_1 - Y_0) - Var(Y_1) - Var(Y_0)}{2\sqrt{Var(Y_0)Var(Y_1)}} \\
&= \frac{E(Y_1 - Y_0)^2 - (E(Y_1) - E(Y_0))^2 - Var(Y_1) - Var(Y_0)}{2\sqrt{Var(Y_0)Var(Y_1)}}
\end{aligned}$$

Since $\varphi(x) = x^2$ is a convex function, the result follows from Lemma 3.1. Sharpness follows from the fact that H_- and H_+ are sharp bounds on the Fréchet Space.$\square$

Proof of Lemma 3.2 : Notice that the absolute value of the program impact can be decomposed as follows

$$\mid y(1) - y(0) \mid = y(1) - y(0) - 2\min(y(1) - y(0), 0)$$

Taking expectations at both sides of the equality, we have

$$m = E(y(1) - y(0)) - 2L$$

From where

$$L = \frac{E(y(1) - y(0)) - m}{2} \tag{31}$$

Since m is positive, the result follows from Lemma 3.1. Alternatively, we can apply Lemma 3.1 directly by noticing that

$$\min(y(1) - y(0), 0) = -\max(y(0) - y(1), 0)$$

and $\varphi(x) = \max(x, 0)$ is a convex function.$\square$

Proof of Proposition 3.3: By Lemma 3.1 and the Frechet bounds, we have

$$\int \psi(y_1 - y_0)dH - \psi(\int y_1 dF_1 - \int y_0 dF_0) \geq \Phi(F_0, F_1)$$

for all $H \in \mathcal{H}(F_1, F_0)$. Define a random variable $Z = Y_1 - Y_0$. By Jensen's inequality

$$\begin{aligned}
\int \psi(z) dH_+ &\geq \psi(\int z dH_+) \\
&= \psi(\int y_1 dH_+ - \int y_0 dH_+) \\
&= \psi(\int y_1 dF_1 - \int y_0 dF_0)
\end{aligned}$$

which is equivalent to $\Phi(F_0, F_1) \geq 0$. The result follows by using the fact that Jensen's inequality holds with equality for the case of strictly convex functions if and only if $Y_1 - Y_0$ is a constant with probability 1.□

Proof of Corollary 3.1: Notice that

$$\begin{aligned}
\Phi_2 &= Var_{H_+}(Y(1) - Y(0)) \\
&= \sigma_1^2 + \sigma_0^2 - 2\rho_{H_+}\sigma_1\sigma_0 \\
&\geq \sigma_1^2 + \sigma_0^2 - 2\sigma_1\sigma_0 \\
&= (\sigma_1 - \sigma_0)^2
\end{aligned}$$

where ρ_{H_+} is the correlation coefficient evaluated at H_+. Hence $\sigma_1 \neq \sigma_0$ implies $\Phi_2 > 0$, and the result follows from Proposition 3.3.□

Proof of Corollary 3.1 when $y(1)$ and $y(0)$ are members of the same location-scale family. Since y_0 and y_1 are members of the same location-scale family, we have

$$y_i = \sigma_i Z + \mu_i$$

for $i = 0, 1$, where $Z \sim f(z)$. Because for any location-scale family it is possible to choose $f(z)$ such that $EZ = 0$ and EZ^2=1, without loss of generality, we choose these values for the first and second moment of Z. Notice that the extreme joint distribution H_+ is obtained when there is maximum correlation between y_1 and y_0.

This occurs when high values of y_1 are "matched" with high values of y_0. This is equivalent to form the pairs $(\sigma_0 z + \mu_0, \sigma_1 z + \mu_1)$ for all z in the support of Z. Hence

$$\begin{aligned}
\int (y(1) - y(0))^2 dH_+ &= \int [(\sigma_1 - \sigma_0)z + (\mu_1 - \mu_0)]^2 df(z) \\
&= E_Z[((\sigma_1 - \sigma_0)^2 Z^2 + 2(\sigma_1 - \sigma_0)(\mu_1 - \mu_0)Z + (\mu_1 - \mu_0)^2] \\
&= (\sigma_1 - \sigma_0)^2 + (\mu_1 - \mu_0)^2
\end{aligned}$$

The result follows from Proposition 3.3.□

Proof of Proposition 3.4: Notice that to test the hypothesis $H_0 : \Phi = 0$ is equivalent to test $H_0^{\varphi} : \varphi(\Phi) = \varphi(0)$. A level $\beta \in (0, 1/2)$ for the latter hypothesis is given by

$$\text{Reject } H_0^{\varphi}\text{: } \varphi(\Phi) = \varphi(0) \text{ if } V_{1-\beta} - \gamma < \hat{\varphi} - \varphi(0)$$

where $V_\beta = F^{-1}(\beta)$. This a straightforward result since under the null we have

$$P(V_{1-\beta} - \gamma < \hat{\varphi} - \varphi(0)) = \beta$$

I will refer to this test as T_1 for the rest of the proof. Let $G(s) = P(\varphi^* < s)$ be the bootstrap cdf of φ^*. Since $\varphi^* = \hat{\varphi} - \gamma + V$, we have

$$\begin{aligned}
G(s) &= P(V < s - \hat{\varphi} + \gamma) \\
&= F_V(s - \hat{\varphi} + \gamma)
\end{aligned}$$

with inverse $G^{-1}(\beta) = F_V^{-1}(\beta) + \hat{\varphi} - \gamma$. I claim that the test T_2 defined as

$$\text{Reject } H_0^{\varphi} \text{ if } \varphi(0) < G^{-1}(F_V(2\gamma + V_\beta))$$

is equivalent to T_1. This is true since

$$\begin{aligned}
G^{-1}(F_V(2\gamma + V_\beta)) &= 2\gamma + V_\beta + \hat{\varphi} - \gamma \\
&= \gamma - V_{1-\beta} + \varphi
\end{aligned}$$

It follows that the test T_3

$$\text{Reject } H_0^{\varphi} \text{ if } \varphi(0) < \min \varphi(\hat{\Phi}^*)$$

is a level β test since, for some $\beta \in (0, 1/2)$

$$\begin{aligned} \varphi(0) &< \min \varphi(\hat{\Phi}^*) \\ &\leq G^{-1}(F_V(2\gamma + V_\beta)) \end{aligned}$$

Finally, let $H(s) = P(\hat{\Psi}^* < s)$ be the bootstrap cdf of $\hat{\Psi}^*$. Since $\varphi(\cdot)$ is a strictly increasing transformation, the quantiles of $\varphi(\hat{\Psi}^*)$ coincide with those of $\hat{\Psi}^*$. Hence, T_3 is equivalent to

$$\text{Reject } H_0 : \Psi = 0 \text{ if } 0 < \min \hat{\Psi}^*$$

since $\min \varphi(\hat{\Psi}^*) = \varphi(\min \hat{\Psi}^*)$. This completes the proof. □

Lemma 6 $\psi(x) = \mid x \mid^\alpha$ *is a strictly convex function for* $\alpha \geq 2$

PROOF: For $\alpha = 2$, the result is immediate since $\psi(x) = x^2$, and $\psi'' > 0$. For $\alpha > 2$, we make use of Pecaric and Dragomir's inequality, which indicates that if $pq(q + p) > 0$, $z_1, z_2 \in \mathbb{R}$, and $\alpha \geq 1$, then

$$\frac{\mid z_1 + z_2 \mid^\alpha}{p + q} \leq \frac{\mid z_1 \mid^\alpha}{p} + \frac{\mid z_2 \mid^\alpha}{q}$$

w.l.g. define $z_1 = \lambda x$, $z_2 = (1 - \lambda)y$, $x, y \in \mathbb{R}$, $p = \lambda$, $q = 1 - \lambda$, and $\lambda \in (0, 1)$. Then we have $\lambda(1 - \lambda) > 0$, and hence

$$\begin{aligned} \mid \lambda x + (1 - \lambda)y \mid^\alpha &\leq \frac{\mid \lambda x \mid^\alpha}{\lambda} + \frac{\mid (1 - \lambda)y \mid^\alpha}{1 - \lambda} \\ &= \lambda^{\alpha - 1} \mid x \mid^\alpha + (1 - \lambda)^{\alpha - 1} \mid y \mid^\alpha \\ &< \lambda \mid x \mid^\alpha + (1 - \lambda) \mid y \mid^\alpha \end{aligned}$$

Where I have used the fact that $\lambda^{\alpha-1} < \lambda$ and $(1-\lambda)^{\alpha-1} < (1-\lambda)$, for $\alpha > 2$.□

Proof of Lemma 3.3: Let $\tau(y_0(q))$ and $\phi(\cdot)$ be defined, respectively, by

$$\tau(y_0(q)) = \inf\{\xi : q \leq F_1(y_0(q) + \xi)\}$$

and

$$\phi(y_0(q)) = F_1^{-1}(F_0(y_0(q)))$$

From the quantile function, we have

$$y_1(q) = F_1^{-1}(q) = \inf\{x : F(y(1) \leq x) > q\}$$

Hence,

$$y_1(q) = \phi(y_0(q)) = F_1^{-1}(q) = \tau(y_0(q)) + y_0(q)$$

The result follows by noticing that $\phi(y_0(q))$ is a non-decreasing function of $y_0(q)$. For a proof of uniqueness see Doksum (1974).□

Proof of Lemma 3.4: Doksum (1974) shows that if $\tau(x)) = \tau$ for x in the support of $y(0)$, then $F_0(x) = F_1(x + \tau)$ for all x. Therefore

$$F_0(y_0(q)) = F_1(y_0(q) + \tau)$$

From the proof of Lemma 3.3, we have

$$y_1(q) = F_1^{-1}(F_0(y_0(q))) = y_0(q) + \tau$$

The result follows.□

Appendix B: Estimation and Bootstrap Algorithm using the Empirical Quantiles

The objective is to estimate bootstrap confidence intervals for the parameters $\theta_- = E_{H_-}[\phi(y_1 - y_0)]$ and $\theta_+ = E_{H_+}[\phi(y_1 - y_0)]$, for some measurable function $\phi(\cdot)$. The data in this problem consists of two independent random samples drawn $Y_i(1) \sim F_1$ and $Y_j(0) \sim F_0$, $i = 1, \ldots, n$, $j = 1, \ldots, m$. Let $\hat{F}_1$ and $\hat{F}_0$ denote the empirical distribution functions implied by these samples.

1. Estimation of θ_- and θ_+

1) Estimate $b = [\gamma \min\{n, m\}]$ empirical quantiles for F_1 and F_0, where $\gamma \in (0, 1)$ and $[\cdot]$ is the integer function. More precisely, for each $t \in \{t_1, \ldots, t_b\}$, $i = 1, 2$, we estimate

$$q_i^{t_j} = \inf\{x : \hat{F}_i(y(i) \leq x) \geq t_j\} \tag{32}$$

2) Let $\hat{Q}_1$ and $\hat{Q}_0$ be the empirical distribution function of the quantiles estimated above, that is to say, a distribution placing a probability mass $\frac{1}{b}$ to each of these quantiles:

$$\hat{Q}_i(x) = \frac{1}{b} \sum_{j=1}^{b} \mathbf{1}(q_i^{t_j} \leq x) \tag{33}$$

3) For all $\mathbf{x} = (x_1, x_2) \in \mathbb{R}^2$, define

$$\hat{H}_-(x_1, x_2) = \max\{\hat{Q}_0(x_1) + \hat{Q}_1(x_2) - 1, 0\} \tag{34}$$

$$\hat{H}_+(x_1, x_2) = \min\{\hat{Q}_0(x_1), \hat{Q}_1(x_2)\} \tag{35}$$

4) Estimate θ using plug-in estimators: $\hat{\theta}_- = \theta(\hat{H}_-)$ and $\hat{\theta}_+ = \theta(\hat{H}_+)$.

2. Estimation of the Extreme distributions H_- and H_+

Define the sequences of quantiles of F_1 and F_0, respectively, by $\{q_1^{t_j}\}$ and $\{q_0^{t_j}\}$.

Let $\mu_i = E_{Q_i}[q_i^t]$ denote the expected value of the chosen quantiles under probability measure Q_i. The correlation coefficient between q_1^t and q_0^t is given by

$$\rho(q_0^t, q_1^t) = \frac{1}{b} \sum (q_1^{t_j} - \mu_1)(q_0^{t_j} - \mu_0) \tag{36}$$

By Lemma 3.5, we know that this coefficient is at its minimum when it is evaluated at H_+, and is at its maximum when evaluated at H_-. We can estimate the extreme distributions H_- and H_+ by applying the following result:

Lemma 7 *(Hardy, Littlewood, and Polya 1952) The sum of products $\sum_i x_i y_i$ is a maximum when both $\{x_i\}$ and $\{y_i\}$ are increasing, and a minimum when one is increasing and the other is decreasing.*

Therefore. by defining $x_j = (q_1^{t_j} - \mu_1)$ and $y_j = (q_1^{t_j} - \mu_0)$, it follows that H_+ is obtained by pairing the largest quantile of F_1 with the largest quantile of F_0, the second largest quantile of F_1 with the second largest quantile of F_0, and so on. To construct H_-, we just need to pair the largest quantile of F_1 with the smallest quantile of F_0, the second largest quantile of F_1, with the second smallest quantile of F_0, and so on.

3. Bootstrap

5) Generate bootstrap random samples from $\hat{F}_1$ and $\hat{F}_0$: $Y_i^*(1) \sim F_1$ and $Y_j^*(0) \sim F_0$, $i = 1, \ldots, n$, $j = 1, \ldots, m$.

Let F_1^* and F_0^* denote the empirical distributions implied by the bootstrap random samples. That is

$$F_i^*(x) = \frac{1}{b} \sum_j \mathbf{1}(y_{ij} \leq x) \tag{37}$$

6) Replicate steps 1)-4) above for the bootstrap distributions F_1^* and F_0^*. That is:

6a) Estimate b empirical quantiles for F_1^* and F_0^*

$$q_i^{t_j *} = \inf\{x : F_i^*(y^*(i) \leq x) \geq t_j\} \tag{38}$$

6b) Let Q_1^* and Q_0^* be the empirical distribution of the quantiles estimated above, that is to say, a distribution placing a probability mass $\frac{1}{b}$ to each of these quantiles.

$$Q_i^*(x) = \frac{1}{b} \sum_{j=1}^{b} \mathbf{1}(q_i^{t_j *} \leq x) \tag{39}$$

6c) Define

$$\begin{aligned} H_-^*(x_1, x_2) &= \max\{Q_0^*(x_1) + Q_1^*(x_2) - 1, 0\} && (40) \\ H_+^*(x_1, x_2) &= \min\{Q_0^*(x_1), Q_1^*(x_2)\} && (41) \end{aligned}$$

6d) Estimate θ_-^* and θ_+^*, respectively, by $\theta_-^* = \theta(H_-^*)$ and $\theta_+^* = \theta(\hat{H}_+^*)$.

7) Repeated independent generation of F_1^* and F_0^* yields a sequence of independent realizations of θ_+^* and θ_-^*, which can be used to approximate their actual bootstrap distribution.

Appendix C: Quantile Treatment Effects

Let $Q_q(Y \mid T)$ be the conditional quantile function of the conditional distribution $F(Y \mid T)$, where $T \in \{0, 1\}$ is the binary variable indicating treatment status: it takes the value of 1 if treated, and 0 otherwise. Assume $F(Y \mid T)$ is continuous and strictly increasing, and that $Q_q(Y \mid T)$ is linear:

$$Q_q(Y \mid T) = \alpha_q + \beta_q T$$

It can be shown that the parameters α_q and β_q can be characterized as follows (Koenker 1978)

$$(\alpha_q, \beta_q) = \arg\min_{(\alpha,\beta)\in\mathbb{R}^2} E[\rho_q(Y - \alpha - \beta T)]$$

where $\rho_q(u) = u(q - I(u < 0))$ is the check function. Let α^* and β^* be the solution to this problem. Then it is easy to show that the QTE can be recovered from here since

$$\tau(q) = y_1(q) - y_0(q) = Q_q(Y \mid T = 1) - Q_q(Y \mid T = 0) = \beta_q^*$$

For the estimation, let $(y_i, T_i)_{i=1}^n$ be a sample from the population. Then we can apply the analog principle and follow Koenker and Bassett (1978) to estimate α and β:

$$(\hat{\alpha}_q, \hat{\beta}_q) = \arg\min_{(\alpha,\beta)\in\mathbb{R}^2} n^{-1} \sum_{i=1}^n \rho_q(Y_i - \alpha - \beta T_i)$$

References

Abadie, A., J. Angrist, and G. Imbens (2002). Instrumental variables estimation of quantile treatment effects. *Econometrica 68*, 399–405.

Andrews, D. (2000). Inconsistency of the bootstrap when a parameter is on the boundary of the parameter space. *Econometrica 70*, 91–117.

Baland, J. M. and J. Robinson (2000). Is child labor inefficient? *Journal of Political Economy 108*(4), 663–679.

Besley, T. and S. Coate (1992). Workfare versus welfare: Incentive arguments for work requirement in poverty-alleviation programs. *American Economic Review 82*(1), 249–261.

Bickel, P., F. Gotze, and W. V. Zwet (1997). Resampling fewer than n observations: gains, losses, and remedies for losses. *Statistica Sinica 7*, 1–31.

Bitler, M., J. Gelbach, and H. Hoynes (2003). What mean impacts miss. Rand Labor and Population Discussion Series.

Brown, M. B. and A. B. Forsythe (1974). Robust tests for equality of variances. *Journal of the American Statistical Association 69*, 364–367.

Das, J., Q. T. Do, and B. Ozler (2004). Conditional cash transfers and the equity-efficiency debate. World Bank paper series, 3280.

Doksum, K. (1974). Empirical probability plots and statistical inference for nonlinear models in the two-sample case. *The annals of statistics 2*, 267–277.

Efron, B. (1979). Bootstrap methods: Another look at the jacknife. *The Annals of Statistics 7*, 1–26.

Efron, B. (1987). Better bootstrap confidence intervals. *Journal of the American Statistical Association 82*, 171–185.

Fields, G. (2001). *Distribution and Development* (First ed.). Russell Sage Foundation-MIT Press.

Frechet, M. (1951). Sur les tableaux de correlation dont les marges sont donnes. *Annals University Lyon: Series A 14*, 53–77.

Hardy, G., J. Littlewood, and G. Polya (1952). *Inequalities* (Second ed.). Cambridge University Press.

Heckman, J. (1978). Dummy endogenous variable in a simultaneous equation system. *Econometrica 46*, 931–961.

Heckman, J., J. Smith, and N. Clements (1997). Making the most out of programme evaluations and social experiments: Accounting for heterogeneity in programme impacts. *The Review of Economic Studies 64*(4), 487–535.

Kanbur, R. and L. Haddad (1997). Are better off households more unequal or less unequal? *Oxford Economic Papers 46*(3), 445–458.

Khandker, Shahidur, and F. Pitt (2003). Subsidy to promote girl's secondary education the female stipend program in bangladesh. Discussion Paper.

Koenker, R. (1978). Regression quantiles. *Econometrica 46*, 33–50.

Levene, H. (1960). Robust tests for equality of variances. In I. Olkin (Ed.), *Contributions to probability and statistics*. Stanford University Press.

Lopez-Acevedo, G. and A. Salinas (2000). Marginal willingness to pay for education and the determinants of enrollment in mexico. World Bank, Working Paper Series, 2405.

Manski, C. (2003). *Partial Identification of Probability Distributions* (First ed.). Springer-Verlag.

Manski, C. F. (1997). Monotone treatment response. *Econometrica 65*, 1311–1334.

Martinelli, C. and S. Parker (2003). Should transfers to poor families be conditional on school attendance? a household bargaining perspective. *International Economic Review 44*(2), 523–544.

Molinari, F. (2005). Missing treatments. Working Paper, Cornell University.

Ruschendorf, L. (1981). Sharpness of frechet-bounds. *Zeitshrift fur Wahrscheinlichkeitstheorie und verwandte Gebiete 57*(4), 293–302.

Tchen, A. (1980). Inequalities for distributions with given marginals. *The Annals of Probability 8*(4), 814–827.

OTHER A.E.M. WORKING PAPERS

WP No	Title	Fee (if applicable)	Author(s)
2006-08	On the Design of an Optimal Transfer Schedule with Time Inconsistent Preferences		Chavez-Martin del Campo, J.
2006-07	Partial Identification of Poverty Measures with Contaminated and Corrupted Data		Chavez-Martin del Campo, J.
2006-06	Beyond Formality and Informality		Guha-Khasnobis, B., Kanbur, R. and E. Ostrom
2006-05	What's Social Policy Got To Do With Economic Growth?		Kanbur, R.
2006-04	Ghana's Economy at Half Century: An Overview of Stability, Growth and Poverty		Aryeetey, E. and R. Kanbur
2006-03	Poverty and Well-Being In Post-Apartheid South Africa: An Overview of Data, Outcomes and Policy		Bhorat, H. and R. Kanbur
2006-02	Epistemology, Normative Theory and Poverty Analysis: Implications for Q-Squared in Practice		Kanbur, R. and P. Shaffer
2006-01	Terrorism and Residential Preferences: Evidence from NYS Polling Data in New York State		Kay, D., Geisler, C. and N. Bills
2005-27	An Econometric Evaluation of a Geopolitical Theory of Oil Price Behavior		Slaibi, A., Chapman, D. and H. Daouk
2005-26	The Politics of Oil and the Aftermath of Civil War in Angola		Kyle, S.
2005-25	A Strategy For Agricultural Development in Angola		S. Kyle
2005-24	New Directions in Development Economics: Theory or Empirics? A Symposium in Economic and Political Weekly		Banerjee, A., Bardhan, P., Basu, K., Kanbur, R., Mookherjee
2005-23	Spatial Inequality and Development: Overview of UNU-WIDER Project		Kanbur, R. and Venables. A.
2005-22	Oil and Politics in Angola		Kyle, S.
2005-21	A Framework for Scaling Up Poverty Reduction, With Illustrations from South Asia		Devarajan, S. and R. Kanbur

Paper copies are being replaced by electronic Portable Document Files (PDFs). To request PDFs of AEM publications, write to (be sure to include your e-mail address): Publications, Department of Applied Economics and Management, Warren Hall, Cornell University, Ithaca, NY 14853-7801. If a fee is indicated, please include a check or money order made payable to Cornell University for the amount of your purchase. Visit our Web site (*http://aem.cornell.edu/research/wp.htm*) for a more complete list of recent bulletins.