WP 2006-07
January 2006

Working Paper

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

Partial Identification of Poverty Measures with Contaminated and Corrupted Data

Juan Carlos Chavez-Martin del Campo

Partial Identification of Poverty Measures with Contaminated and Corrupted Data

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

January 12, 2006

Abstract

This paper applies a partial identification approach to poverty measurement when data errors are non-classical in the sense that it is not assumed that the error is statistically independent of the outcome of interest, and the error distribution has a mass point at zero. This paper shows that it is possible to find non-parametric bounds for the class of additively separable poverty measures. A methodology to draw statistical inference on partially identified parameters is extended and applied to the setting of poverty measurement. The methodology developed in this essay is applied to the estimation of poverty treatment effects of an anti-poverty program in the presence of contaminated data.

JEL Classification: C14, I32.
Keywords: *Poverty Measurement, Partial Identification, Contamination Model, Statistical Inference, Nonparametric bounds, Poverty Comparisons.*

*I am greatly indebted to Francesca Molinari for her continued guidance and feedback. I also thank Martin Ravallion, participants at the 2004 Latin American Meeting of the Econometric Society in Santiago de Chile, the 2004 Northeast Universities Development Consortium Conference at HEC Montreal, and the 2005 Econometric Society World Congress at University College London for helpful comments and suggestions. 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

Much of the statistical analysis of poverty measurement regards the data employed to estimate a poverty measure as error-free. However, it is amply recognized that measurement error is a very common phenomenon for most data sets used in the estimation of poverty. This problem is particularly relevant for developing countries, where the majority of the poor are concentrated, since financial, technological, and logistical constraints are more likely to affect the quality of the data.

Measurement error can affect the estimation of poverty in different ways. For example, the poverty line may be set for heterogenous groups of people without considering idiosyncratic differences in the cost of basic needs, arbitrary imputations may be made when missing and zero outcomes appear in the sample, and the variable of interest may be misreported by an important subset of survey respondents. Often the methodologies applied to solve these problems are arbitrary; at the same time, the results are highly sensitive to such adjustments. For instance, Szekely, Lustig, Cumpa and Mejia (2000) applied several techniques to adjust for misreporting in Latin America. In the case of Mexico they found that, depending on the method for performing the adjustment, either 14 percent or 76.6 percent of the population is below the poverty line (in absolute terms it implies a difference of 57 million individuals). This has important policy implications since, depending on which of these numbers is used as a reference, the amount of resources directed to social programs can be considered either appropriate or totally insufficient.

Several approaches have been developed in order to analyze the effects of measurement error on poverty measurement. For instance, Chesher and Schluter (2002) study multiplicative measurement error distributed continuously and independently of true income to investigate the sensitivity of welfare measures to alternative amounts of measurement error. Ravallion (1994) considers additive random errors when estimating individual-specific poverty lines, finding that heterogeneity in error distributions generates ambiguous poverty rankings. An alternative approach, robust estimation, aims at developing point estimators that are not highly sensitive to errors in the data.[1] The objective is to guard against worse-case scenarios that errors in the data could conceivably produce. In that sense it takes an ex-ante perspective of the problem. Cowell and Victoria-Feser (1996) apply this approach to poverty measurement by using the concept of the influence function to assess the influence of an infinitesimal amount of contamination upon the value of a poverty statistic (Hampel 1974). They find that poverty measures that take as their primitive concept poverty gaps rather incomes of the poor are in general robust under this criterion.

In the present study, we do not consider classical measurement error, that is to say, we do not assume the existence of chronic errors affecting every observation, neither do we assume that the outcome of interest is statistically independent of the error. Instead of assuming that the error distributions have no mass point at zero, we consider the impact of intermittent errors by setting an upper bound to the proportion of gross errors within the data. Since a poverty measure is not point identified under the assumptions of the model of errors under consideration, we follow Horowitz and Manski (1995) and apply a partial identification approach.[2] By using a fully non-parametric method, we show that for the family of additively separable poverty measures it is possible to find identification regions under very mild assumptions.

The paper is organized as follows. Section 2 introduces some important concepts for poverty measurement. Section 3 states the problem formally, presenting both the contaminated and corrupted sampling models within the context of poverty measurement. Section 4 investigates the identification region for additively separable poverty measures (ASP). It is shown that we can find upper and lower bounds for this class of poverty measures with both contaminated and corrupted data. Section 5 character-

[1] See Hampel et al (1986) and Huber (1981) for a comprehensive treatment of robust inference.

[2] Examples of applications of this approach in other settings are Molinari (2005a) and Dominitz and Sherman (2005). See Manski (2003) for an overview of this literature

izes the identification regions of ASP measures through their length and breakdown points. Section 6 applies two conceptually different types of confidence intervals for partially identified poverty measures. The implications for hypothesis testing when a poverty measure is not point identified are also discussed. Section 7 provides some insight on the effect of both data contamination and data corruption for poverty comparisons. Sections 8 and 9 give two empirical illustrations of the methodology developed in the paper. Most of the mathematical details are in the Appendix.

2 Poverty Measurement: Conceptual Framework

Let $\mathcal{A}$ denote the $\sigma-$algebra of Lebesgue measurable sets on $\mathbb{R}$. Let $\mathcal{P}$ denote the set of all probability distributions on $(\mathbb{R}, \mathcal{A})$. Thus for any $P \in \mathcal{P}$ the triple $(\mathbb{R}, \mathcal{A}, P)$ is a probability space. Let $z \in \mathbb{R}_{++}$ be the poverty line.

A person is said to be in poverty if her income, $y \in \mathbb{R}$, or any other measure of her economic status is strictly below z. An aggregate poverty index is defined as a functional of the distribution $P \in \mathcal{P}$. Formally:

Definition 1 *A Poverty Index is a functional* $\Pi(P; z) : \mathcal{P} \times \mathbb{R}_{++} \rightarrow \mathbb{R}$ *that indicates the degree of poverty when a particular variable has distribution* P *and the poverty line is* z.

An important type of poverty measures is the *Additively Separable Poverty*(ASP) class which is defined as follows:

$$\Pi(P; z) = \int \pi(y; z) dP \tag{1}$$

where $\pi(y; z) : \mathbb{R} \times \mathbb{R}_{++} \rightarrow \mathbb{R}$ is the poverty evaluation function, an indicator of the severity of poverty for a person with income y when the poverty line is fixed at z.

Since the axiomatic approach to poverty measurement proposed by Sen (1976), most economists interested in the phenomenon of poverty have quantified poverty in a manner consistent with those principles. One of those principles, the *focus axiom*, requires a poverty measure to be independent of the income distribution of the non poor. The *monotonicity axiom* says that, everything else equal, a reduction in the income of a poor individual must increase the poverty measure; the *transfer axiom* emphasizes the positive effect of a regressive transfer on the poverty measure, that is to say, given other things, a pure transfer of income from a poor individual to any other individual that is richer must increase the poverty measure. Finally, Kakwani (1980) has proposed a 4th property that prioritizes transfers taking place down in the distribution, other things being equal. These distributional concerns are made operational through the characteristics of the poverty evaluation function $\pi(y; z)$. It is usually assumed that $\pi(y; z)$ is continuous for $y < z$, non increasing in its first argument and non decreasing in its second argument. It is also assumed that $\pi(y; z)$ is convex in its first argument and $\pi(y; z) = 0$ for $y \geq z$.

2.1 Specific Poverty Measures

Watts (1968) proposed a poverty measure which is defined as follows:

$$\Pi_W = \int \mathbf{1}(y < z) \ln(\frac{y}{z}) dP \tag{2}$$

This poverty measure satisfies Sen's monotonicity and transfer axioms as well as Kakwani's transfer-sensitivity axiom.

Foster, Greer and Thorbecke (1984) proposed an $\alpha-$class of poverty measures, Π_α, which can be obtained by:

$$\Pi_\alpha = \int \mathbf{1}(y < z)(1 - \frac{y}{z})^\alpha dP \tag{3}$$

Π_α satisfies monotonicity axiom for $\alpha > 0$, transfer axiom for $\alpha > 1$, and transfer sensitivity axiom for $\alpha > 2$.

Hagenaars (1987) provided a poverty measure that satisfies all three axioms. The specific poverty measure he gave is

$$\Pi_H = \int \mathbf{1}(y < z)(1 - \frac{\ln y}{\ln z})dP \tag{4}$$

Finally, we consider the Clark, Hemming and Ulph (1981) poverty measure:

$$\Pi_\beta = \frac{1}{\beta}\int \mathbf{1}(y < z)(1 - (\frac{y}{z})^\beta)dP \tag{5}$$

which satisfies the monotonicity axiom for all $\beta > 0$, and both transfer axioms for $\beta < 1$. Finally, Chakravarty (1983) derived a poverty measure which is equal to $\Pi_{Ch} = \beta\Pi_\beta$. This measure also satisfies all three axioms for $\beta \in (0, 1)$.

3 Statement of the Problem

Let each member j of population J be characterized by the tuple (y_1^j, y_0^j) in the space $\mathbb{R}\times\mathbb{R}$, where y_1^j is the outcome of interest denoting the "true" equivalent income (or expenditure) for a given poverty line z. Let the random variable $(y_1, y_0) : J \longrightarrow \mathbb{R} \times \mathbb{R}$ have distribution $P(y_1, y_0)$. Let a random sample be drawn from $P(y_1, y_0)$. Let's assume that instead of observing y_1, one observes a random variable y defined by:

$$y \equiv wy_1 + (1 - w)y_0 \tag{6}$$

Realizations of y with $w = 0$ are said to be data errors, those with $w = 1$ are error-free, and y itself is a contaminated version of y_1. Let $Q(y)$ denote the distribution of the observable y. Let $P_i = P_i(y_i)$ denote the marginal distribution of y_i. Let $P_{ij} = P_{ij}(y_i \mid w = j)$ denote the distribution of y conditional on the event $w = j$ for

$i, j \in \{0, 1\}$. Let $p = P(w = 0)$ be the marginal probability of a data error. With data errors, the sampling process does not identify P_1 (the object of interest) but only $Q(y)$, the distribution of the observable y. By the law of total probability, these two distributions can be decomposed as follows:

$$P_1 = (1 - p)P_{11} + pP_{10} \tag{7}$$

$$Q(y) = (1 - p)P_{11} + pP_{00} \tag{8}$$

This problem can be approached from different perspectives. In robust estimation P_1 is held fixed and $Q(y)$ is allowed to range over all distributions consistent with both equations. In the context of poverty measurement, the objective would be to estimate the maximum possible distance between $\Pi(Q; z)$ and $\Pi(P_1; z)$. In contrast, the present analysis holds $Q(y)$ fixed because it is identified by the data, and P_1 is allowed to range over all distributions consistent with (3) and (4). This approach recognizes that the parameter of interest might not be point identified, but it can often be bounded.

The sampling process reveals only the distribution $Q(y)$. However, informative identification regions emerge if knowledge of the empirical distribution is combined with a non-trivial upper bound, λ, on p.

This investigations analyzes two different cases of data errors. In the first case, we will assume that the occurrence of data errors is independent of the sample realizations from the population of interest. Formally

$$P_1 = P_{11} \tag{9}$$

This particular model of data errors is known as "contaminated data" or "contaminated sampling" model (Huber 1981). In the other case, (9) does not hold and it is only assumed that there exists a non-trivial upper bound on the error probability.

Horowitz and Manski (1995) refer to this case as "corrupted sampling".

Define the sets

$$\mathcal{P}_1(p) \equiv \mathcal{P} \cap \{(1-p)\phi_{11} + p\phi_{10} : (\phi_{11}, \phi_{10}) \in \mathcal{P}_{11}(p) \times \mathcal{P}\} \tag{10}$$

$$\mathcal{P}_{11}(p) \equiv \mathcal{P} \cap \left\{ \frac{Q - p\phi_{00}}{1-p} : \phi_{00} \in \mathcal{P} \right\} \tag{11}$$

If there exists a non-trivial upper bound, λ, on the probability of data errors, then it can be proved that P_{11} and P_1 belong to the sets $\mathcal{P}_{11}(\lambda)$ and $\mathcal{P}_1(\lambda)$ respectively, where $\mathcal{P}_{11}(\lambda) \subset \mathcal{P}_1(\lambda)$. These restrictions are sharp in the sense that they exhaust all the available information, given the maintained assumptions (Horowitz and Manski 1995).

4 Partial Identification of Poverty Measures

Suppose that a proportion $p < 1$ of the data is erroneous. Furthermore, assume there exists a non-trivial upper bound, λ, on p: $p \leq \lambda < 1$.[3] From the analysis above, we know that the distribution of interest P_1 is not identified: i.e. $\mathcal{P}_1(\lambda)$ is not a singleton.

Even though P_1 is not point identified, it is partially identified in the sense that it belongs to the identification region $\mathcal{P}_1(\lambda)$. There is a mapping from this set into the domain in $\mathbb{R}$ of a given poverty measure Π. Therefore, the question arises whether there is a way to characterize the identification region of Π. As we will see below, it is possible to do so for the class of ASP poverty measures by ordering the distributions in $\mathcal{P}_\lambda$ according to a stochastic dominance criterion. Such criterion is defined as

[3] In practice, upper bounds on the probability of data errors can be estimated from a validation data set or by the proportion of imputed data in the sample. See Kreider and Pepper (2004) for an application of a validation model.

follows:

Definition 2 *Let $F, G \in \mathcal{P}$. Distribution F First Order Stochastically dominates (FOD) distribution G if*

$$F((-\infty, x]) \leq G((-\infty, x])$$

for all $x \in \mathbb{R}$.

In the case of monotone functions, there is a well-known result that will be helpful to obtain identification regions for the ASP measures:

Lemma 1 *The Distribution F first-order stochastically dominates the distribution G if and only if, for every non decreasing function $\varphi : \mathbb{R} \to \mathbb{R}$, we have*

$$\int \varphi(x) dF(x) \geq \int \varphi(x) dG(x) \tag{12}$$

Finally, let me introduce a basic concept that is a building block for identification regions.

Definition 3 *For $\alpha \in (0, 1]$, the α-quantile of $Q(y)$ is given by*

$$r(\alpha) = inf\{t : Q((-\infty, t]) \geq \alpha\}.$$

Now we can state the main result of this section. Following the approach of Horowitz and Manski (1995) to find sharp bounds on parameters that respect stochastic dominance [4] we can construct identification regions for ASP measures.

Proposition 1 *Let it be known that $p \leq \lambda < 1$. Define probability distributions L_λ and U_λ on $\mathbb{R}$ as follows:*

$$L_\lambda = \begin{cases} \frac{Q(y \leq t)}{1-\lambda} & for\ t < r(1-\lambda) \\ 1 & otherwise \end{cases}$$

[4] A parameter $\delta(\cdot)$ respects stochastic dominance if $\delta(F) \geq \delta(G)$ whenever F FOD G.

$$U_\lambda = \begin{cases} 0 & \textit{for } t < r(\lambda) \\ \frac{Q(y \leq t) - \lambda}{1-\lambda} & \textit{otherwise} \end{cases}$$

If $\Pi(P;z)$ belongs to the family of Additively Separable Poverty Measures and the poverty evaluation function is non-increasing in y, then identification regions for $\Pi(P_{11};z)$ and $\Pi(P_1;z)$ are given by:

$$\mathbf{H}[\Pi(P_{11};z)] = [\Pi_l(U_\lambda;z), \Pi_u(L_\lambda;z)] \tag{13}$$

and

$$\mathbf{H}[\Pi(P_1;z)] = [(1-\lambda)\Pi_l(U_\lambda;z) + \lambda\psi_0, (1-\lambda)\Pi_u(L_\lambda;z) + \lambda\psi_1] \tag{14}$$

where $\psi_0 = \inf_{y\in\mathbb{R}_+} \pi(y;z)$ and $\psi_1 = \sup_{y\in\mathbb{R}_+} \pi(y;z)$.

PROOF: See Appendix.

These results are quite intuitive. In the case of contaminated data, the smallest feasible value of $\Pi(P_{11};z)$ occurs when we place all of the erroneous data as far out as possible in the left-hand tail of the observed distribution Q. Similarly, to obtain the largest feasible value of $\Pi(P_{11};z)$, L_λ places all of the erroneous data as far out as possible in the right-hand tail of the observed income distribution. If the data is corrupted, we follow a similar procedure, placing all of the erroneous data at $\inf_y \pi(y;z)$ and $\sup_y \pi(y;z)$ instead.

Example 1 *Assume $P_1 = P_{11}$. Let $Q(y) = U[0,1]$, $0 < p < \lambda < z < 1-\lambda$. Let the poverty measure be given by $\varphi = \int_0^\infty 1(y<z)d\phi$. Then, $\varphi(P_1;z) \in [\frac{z-\lambda}{1-\lambda}, \frac{z}{1-\lambda}]$. If $P_1 \neq P_{11}$ then $\varphi(P_1;z) \in [z-\lambda, z+\lambda]$. Notice that $\varphi(Q;z)$ belongs to both intervals.*

5 Characterizing Identification Regions

The objective of this section is to describe the properties of the identification region for ASP measures. Our approach is not normative in that we are not arguing that one poverty measure is better than another based on our findings. We analyze identification regions through two concepts: identification breakdown points and length of the identification region, with the hope of shedding some light on the identification properties of poverty measures.

5.1 Identification Breakdown Points for Poverty Measures

We denote by $\mathcal{D}$ the family of ASP measures indexed by j with poverty evaluation function $\pi_j(y;z)$ satisfying $\pi_j(y;z) = 0$ for all $y \geq z$, increasing in its second argument, decreasing in its first argument, and continuous and convex for all $y < z$. Moreover, we assume the existence of a constant $c_j \in \mathbb{R}_+$ such that $\pi_j(0;z) = c_j$. We denote by $\mathfrak{R}_j = \{\Pi_j(P;z); P \in \mathcal{P}\}$ the range of a poverty measure $\Pi_j(P;z)$ in $\mathcal{D}$. More precisely, the range of a poverty measure in $\mathcal{D}$ is given by $\mathfrak{R}_j = [0, c_j]$.[5]

From the literature on robust estimation we borrow the concept of breakdown point which in the present setting can be interpreted as *the largest fraction of erroneous data that can be in a sample without driving a poverty measure to either boundary of its range.* However, as noticed by Horowitz and Manski (1995), there are some conceptual differences between the breakdown point in robust estimation and its counterpart in identification analysis. While in the partial identification approach λ is evaluated at the empirical distribution Q, in robust estimation it is evaluated at the distribution of interest P_1. More formally, *the identification breakdown point of a poverty measure* $\Pi(P;z)$ when data are contaminated can be constructed as follows: for some ASP measure in $\mathcal{D}$ define

$$\phi(\lambda_j) = \Pi_j(L_\lambda; z) - c_j \quad (15)$$

$$\psi(\lambda_j) = \Pi_j(U_\lambda; z) \quad (16)$$

[5]Most of the ASP measures used in empirical work belong to this class. For example, the Foster, Greer and Thorbecke (1984), and the Clark, Hemming and Ulph (1981) families of poverty measures are two elements of $\mathcal{D}$.

and let $\lambda_j^\phi = \sup\{\lambda : \phi_j(\lambda) < 0\}$, and $\lambda_j^\psi = \sup\{\lambda : \psi_j(\lambda) > 0\}$. The identification breakdown point for an ASP measure is given by:

$$\lambda_j^* = \min\{\lambda_j^\phi, \lambda_j^\psi\} \tag{17}$$

Let $H_Q = \int 1(y < z)dQ$ be the head-count ratio or proportion of the poor for the observed distribution Q, and let λ^H be its breakdown point. Clearly the head-count ratio is an element of $\mathcal{D}$. We have the following proposition:

Proposition 2 *For all $Q \in \mathcal{P}$, we have*

$$\lambda^H = \inf\{\lambda_j^* : j \in \mathcal{D}\}$$

PROOF: See Appendix.

Therefore, the breakdown point for the head-count ratio is a lower-bound of the set $\mathcal{D}$.

5.2 Length

Another way to "compare" the different poverty measures is through the length of their identification regions. Although we are not arguing here that one can choose one poverty measure over another based on this criterion, the results obtained in this section provide some initial insights about the behavior of the different poverty measures for the model of errors under consideration. To formalize the analysis, let $m : \mathcal{B} \to \mathbb{R}_+$ be the Lebesgue measure on the Borel sets, $\mathcal{B}$, of $\mathbb{R}_+$. Here is the main result of this section:

Proposition 3 *Let $\Pi_1(P;z) : \mathcal{P} \times \mathbb{R}_{++} \to \mathbb{R}$ and $\Pi_2(P;z) : \mathcal{P} \times \mathbb{R}_{++} \to \mathbb{R}$ be two additively separable poverty measures with non-increasing evaluation functions $\pi_1(y;z) : \mathbb{R}_{++} \times \mathbb{R} \to \mathbb{R}$ and $\pi_2(y;z) : \mathbb{R}_{++} \times \mathbb{R} \to \mathbb{R}$, respectively. Sup-*

pose that $\pi_2(y;z) \geq \pi_1(y;z)$ *for all* $y < z$ *and* $\pi_2(y;z) = \pi_1(y;z)$*, otherwise. Let* $z \leq \max\{r(\lambda), r(1-\lambda)\}$ *If the data is either corrupted or contaminated, then*

$$m(\mathbf{H}[\Pi_2]) \geq m(\mathbf{H}[\Pi_1])$$

PROOF: See Appendix.

We can get a similar result by imposing more assumptions on the "shape" of the poverty evaluation function. In particular, we can use the fact that some families of poverty measures are generated by "convexifying" a poverty evaluation function in order to show the existence of length orderings within families of poverty measures. The following two corollaries state this result more formally:

Corollary 1 *Let* $\pi_1(y;z) : \mathbb{R}_+ \times \mathbb{R}_{++} \to \mathbb{R}$ *be a non increasing, and continuous on* $(0,z)$ *poverty evaluation function, with* $z \leq \max\{r(\lambda), r(1-\lambda)\}$*, and* f *be a convex function on* $\pi_1(y;z)$ *such that*

A1. $\pi_2(y;z) = f \circ \pi_1(y;z)$

A2. $f(\pi_1(0;z)) \leq \pi_1(0;z)$

A3. $f(\pi_1(z;z)) = \pi_1(z;z)$

Then $m(\mathbf{H}[\Pi_2]) \leq m(\mathbf{H}[\Pi_1])$.

PROOF: See Appendix.

Corollary 2 *Given two continuous poverty evaluation functions* $\pi_1(y;z) : \mathbb{R}_+ \times \mathbb{R}_{++} \to \mathbb{R}$ *and* $\pi_1(y;z) : \mathbb{R}_+ \times \mathbb{R}_{++} \to \mathbb{R}$*,* $z \leq \max\{r(\lambda), r(1-\lambda)\}$ *such that*

A4. $\pi_1(0;z) = \pi_2(0;z)$

A5. $\pi_1(y;z) = \pi_2(y;z)$*, for all* $y \geq z$

A6. $\pi_i' < 0$*,* $\pi_i'' > 0$ *on* $(0,z)$*,* $i = 1,2$

A7. $-\frac{\pi_1(y;z)''}{\pi_1(y;z)'} \geq -\frac{\pi_2(y;z)''}{\pi_2(y;z)'}$ *uniformly on* $(0,z)$

Then $m(\mathbf{H}[\Pi_2]) \leq m(\mathbf{H}[\Pi_1])$.

PROOF: See Appendix.

Example 2 *An α-ordering.*

Let $\pi_\alpha(y;z) : \mathbb{R}_+ \times \mathbb{R}_{++} \to \mathbb{R}$ be defined as follows:

$$\pi_\alpha = \begin{cases} (1 - \frac{y}{z})^\alpha & \text{if } y \in [0, z) \\ 0 & \text{if } y > z \end{cases}$$

Define $f_c(x) = x^c$. Clearly this is a convex function for all $x > 0$ and $c \geq 1$. Take any positive integers α_1 and α_2 such that $\alpha_1 = k\alpha_2 > 0$. Hence,

$$\pi_{\alpha_1}(y;z) = f_k \circ \pi_{\alpha_2}$$

By Corollary 1.1, $m_{\alpha_2} \geq m_{\alpha_1}$.

Example 3 *An e-ordering*

Let $\pi_{Ch}(y;z) : \mathbb{R}_+ \times \mathbb{R}_{++} \to \mathbb{R}$ be defined as follows:

$$\pi_{Ch} = \begin{cases} 1 - (\frac{y}{z})^e & \text{if } y \in [0, z) \\ 0 & \text{if } y > z \end{cases}$$

for $e \in (0,1)$. After some algebraic manipulations we have:

$$-\frac{\pi''_{Ch}}{\pi'_{Ch}} = \frac{(1-e)}{y}$$

on $(0,z)$. Therefore, $m_e(\mathbf{H}[\Pi_{Ch}])$ is decreasing on e by Corollary 1.2.

Example 4 *Length rankings.*

Let $\alpha \geq 1$, $\beta \in (0,1)$, $e \in (0,1)$. Then it is easy to show, applying Proposition 1.3, that the following length rankings hold: $m(\mathbf{H}[\Pi_W]) \geq m(\mathbf{H}[\Pi_C]) \geq m(\mathbf{H}[\Pi_\alpha])$ and $m(\mathbf{H}[\Pi_W]) \geq m(\mathbf{H}[\Pi_C]) \geq m(\mathbf{H}[\Pi_{Ch}])$.

6 Statistical Inference for Partially Identified Poverty Measures

In this section, we obtain two conceptually different types of confidence sets for the identification regions of poverty measures. The first type of confidence set uses the Bonferroni's inequality to develop confidence intervals that asymptotically cover the entire identification region with at least probability γ. For the second type of confidence set, we follow Imbens and Manski (2004) by applying confidence intervals that asymptotically cover the true value of the poverty measure with at least this probability. We also discuss some implications of this methodology for hypothesis testing in the context of partially identified poverty measures.

6.1 Confidence Intervals

Let $(\mathbb{R}, \mathcal{A}, Q)$ be a probability space, and let $\mathcal{P}$ be a space of probability distributions. The distribution Q is not known, but a random sample $y_1, y_2, \ldots, y_n$ is available.

In the point identified case ($\lambda = 0$), a consistent estimator of the class of ASP measures is given by

$$\hat{\Pi} = \frac{1}{n}\sum_{i=1}^{n} \pi(y_i; z) \tag{18}$$

where $\pi(y; z)$ is a measurable poverty evaluation function. By applying The Central Limit Theorem, the standard $100 \cdot \gamma\%$ confidence interval for $\Pi(P; z)$ is given by:

$$CI_{\gamma}^{\Pi} = \left[\hat{\Pi} - z_{\frac{\gamma+1}{2}}\frac{\hat{\sigma}}{\sqrt{n}}, \hat{\Pi} + z_{\frac{\gamma+1}{2}}\frac{\hat{\sigma}}{\sqrt{n}}\right] \tag{19}$$

where $\hat{\sigma} = \sigma + o_p(1)$ and z_{τ} is the τ quantile of the standard normal distribution.[6]

To derive the asymptotic properties for the Bonferroni confidence set, we will

[6]Kakwani (1993) describes this methodology for ASP measures.

make use of a result on L-statistics due to Stigler (1973), who explores the asymptotic behavior of trimmed means. Define the confidence interval $CI_{\gamma}^{[\Pi_l,\Pi_u]}$ as

$$CI_{\gamma}^{[\Pi_l,\Pi_u]} = \left[\hat{\Pi}_l - z_{\frac{\gamma+1}{2}}\frac{\hat{\sigma}_l}{\sqrt{n}}, \hat{\Pi}_u + z_{\frac{\gamma+1}{2}}\frac{\hat{\sigma}_u}{\sqrt{n}}\right] \tag{20}$$

Where $\hat{\Pi}_l$, $\hat{\Pi}_u$, $\hat{\sigma}_l^2$, and $\hat{\sigma}_u^2$ are estimators satisfying, respectively

A8. $\hat{\Pi}_l = \Pi_l + o_p(1)$

A9. $\hat{\Pi}_u = \Pi_l + o_p(1)$

A10. $\hat{\sigma}_l^2 = \frac{Var_{U_\lambda}(\pi(y;z))+(\pi(r(1-\lambda))-\Pi_l)\lambda}{1-\lambda} + o_p(1)$

A11. $\hat{\sigma}_u^2 = \frac{Var_{L_\lambda}(\pi(y;z))+(\pi(r(\lambda))-\Pi_u)\lambda}{1-\lambda} + o_p(1)$

We have the following result

Proposition 4 *Let $\lambda < 1$ be known. Assume $\int \pi(y;z)^2 dQ < \infty$. Let $r(1-\lambda)$ and $r(\lambda)$ be continuity points of $Q(y)$. Let the poverty evaluation function, $\pi(y;z)$, be a non-increasing function that is continuous at $r(1-\lambda)$ and $r(\lambda)$. Then*

$$\lim_{n\to\infty} \mathbb{P}([\Pi_l, \Pi_u] \subset CI_{\gamma}^{[\Pi_l,\Pi_u]}) \geq \gamma \tag{21}$$

PROOF: See Appendix.

For the second type of confidence interval, define $\Delta = \Pi_U - \Pi_L$ and $\hat{\Delta} = \hat{\Pi}_U - \hat{\Pi}_L$ and consider the following set of regularity conditions, which are equivalent to the assumptions imposed by Imbens and Manski (2004).[7]

A13. $Q(y) \in \mathcal{F}$, where $\mathcal{F}$ is the set of distribution functions for which $\int \mid \pi(y;z) \mid^3 dQ < \infty$, Q'' is bounded in the neighborhoods of $r(\lambda)$ and $r(1-\lambda)$ while $Q'(r(\lambda)) > 0$ and $Q'(r(1-\lambda)) > 0$.

A14. $\underline{\sigma}^2 \leq \sigma_l^2, \sigma_u^2 \leq \overline{\sigma}^2$ for some positive and finite $\underline{\sigma}^2$ and $\overline{\sigma}^2$.

A15. $\Pi_u - \Pi_l \leq \overline{\Delta} < \infty$

[7]More precisely, we have made use of the results on uniform convergence of trimmed means developed by De Wett (1976) to develop a set of regularity conditions equivalent to those required by Imbens and Manski (2004) to obtain their asymptotic result.

A16. For all $\epsilon > 0$ there are $\nu > 0$, K and n_0 such that $n \geq n_0$ implies $Pr\left(\sqrt{n} \mid \hat{\Delta} - \Delta \mid > K\Delta^{\nu}\right) < \epsilon$, uniformly in $Q \in \mathcal{F}$.

Define the confidence interval $\overline{CI}_{\gamma}^{\Pi}$ as:

$$\overline{CI}_{\gamma}^{\Pi} = \left[\widehat{\Pi}_l - \frac{\overline{C}_n \hat{\sigma}_l}{\sqrt{n}}, \widehat{\Pi}_u + \frac{\overline{C}_n \hat{\sigma}_u}{\sqrt{n}}\right] \tag{22}$$

where $\overline{C}_n$ satisfies

$$\Phi\left(\overline{C}_n + \sqrt{n}\frac{\hat{\Delta}}{max(\hat{\sigma}_l, \hat{\sigma}_u)}\right) - \Phi\left(-\overline{C}_n\right) = \gamma \tag{23}$$

Proposition 5 *Let $\lambda < 1$. Let $r(1-\lambda)$ and $r(\lambda)$ be continuity points of $Q(y)$. Let the poverty evaluation function, $\pi(y; z)$, be a non-increasing function that is continuous at $r(1-\lambda)$ and $r(\lambda)$. Suppose **A13-A16** hold. Then*

$$lim_{n\to\infty} inf_{P\in\mathcal{P}} \mathbb{P}\left(\Pi \in \overline{CI}_{\gamma}^{\Pi}\right) \geq \gamma \tag{24}$$

PROOF: See Appendix.

6.2 Hypothesis Testing

Consider the implications of testing hypothesis of the form:

$$H_0 : \Pi = \Pi_0$$

versus

$$H_1 : \Pi \neq \Pi_0$$

When a parameter is not point identified, the power of a test is not a straightforward extension of the point identified case. For instance, consider the test

reject H_0 if $\frac{\sqrt{n}(\hat{\Pi}_l - \Pi_0)}{\sigma_l} > z_{\frac{\gamma+1}{2}}$ or $\frac{\sqrt{n}(\hat{\Pi}_u - \Pi_0)}{\sigma_u} < -z_{\frac{\gamma+1}{2}}$

The rejection region is

$$R = \{(y_1, \dots, y_n) : \frac{\sqrt{n}(\hat{\Pi}_l - \Pi_0)}{\sigma_l} < z_{\frac{\gamma+1}{2}} \text{ or } \frac{\sqrt{n}(\hat{\Pi}_u - \Pi_0)}{\sigma_u} < -z_{\frac{\gamma+1}{2}}\}$$

and the power function is defined by

$$\beta_n(\Pi) = \mathbb{P}_\Pi((y_1, \dots, y_n) \in R)$$

Define the events

$$A_n = \{Y_n + \frac{\sqrt{n}(\Pi_l - \Pi_0)}{\hat{\sigma}_l} > z_{\frac{\gamma+1}{2}}\}$$

$$B_n = \{Z_n + \frac{\sqrt{n}(\Pi_u - \Pi_0)}{\hat{\sigma}_u} < -z_{\frac{\gamma+1}{2}}\}$$

where $Y_n = \frac{\sqrt{n}(\hat{\Pi}_l - \Pi_l)}{\hat{\sigma}_l}$ and $Z_n = \frac{\sqrt{n}(\hat{\Pi}_u - \Pi_u)}{\hat{\sigma}_u}$. From proposition 1.4, we can deduce that this test has a level $1 - \gamma$ since

$$\begin{aligned} \lim_{n\to\infty} \beta_n(\Pi_0) &= \lim_{n\to\infty} \mathbb{P}(A_n \cup B_n) \\ &\leq 1 - \lim_{n\to\infty} \mathbb{P}([\Pi_l, \Pi_u] \subset CI_\gamma^{[\Pi_l,\Pi_u]}) \\ &\leq 1 - \gamma \end{aligned}$$

Next, suppose the true value of Π is $\Pi^* \neq \Pi_0$. If Π is point identified, the probability of correctly rejecting the null hypothesis, H_0, tends to 1 asymptotically. On the other hand, if the parameter is not point identified, the power of the test for values other than Π_0 is not longer equal to one in general. To verify that this is the case, it will be helpful to divide the analysis in several cases:

i) $\Pi_0 \in [\Pi_l, \Pi_u]$

In this case $\lim_{n\to\infty} \beta_n(\Pi^*) = \lim_{n\to\infty} \beta_n(\Pi_0) \leq 1 - \gamma$. Hence, a type II error is more likely to arise whenever Π_0 belongs to the identification region.

ii) $\Pi_0 < \Pi_l$

Notice that

$$\begin{aligned}\lim_{n\to\infty}\beta_n(\Pi^*) &\geq \lim_{n\to\infty}\mathbb{P}(A_n\cup B_n)\\ &\leq \lim_{n\to\infty}\mathbb{P}(\frac{\sqrt{n}(\hat{\Pi}_l-\Pi_l)}{\hat{\sigma}_l}+\frac{\sqrt{n}(\Pi_l-\Pi_0)}{\hat{\sigma}_l})\\ &= 1\end{aligned}$$

where I have used the fact that $\frac{\sqrt{n}(\hat{\Pi}_l-\Pi_l)}{\hat{\sigma}_l}+\frac{\sqrt{n}(\Pi_l-\Pi_0)}{\hat{\sigma}_l}$ will converge to $+\infty$ in probability. Since $\beta(\Pi^*)$ is a probability measure, we have $\lim_{n\to\infty}\beta_n(\Pi^*)=1$.

iii) $\Pi_0>\Pi_u$

By a similar argument to the one applied in *ii)*, we have $\lim_{n\to\infty}\beta_n(\Pi^*)=1$.

Interestingly, the power $\beta(\Pi^*)$ is a decreasing function of λ because the size of the identification region is positively related to it: the larger the value of the upper bound λ, the more likely it is that Π_0 belongs to the identification region, implying a higher a probability that a type II error will occur.

7 Poverty Comparisons

This section addresses both identification and inference problems when comparing some poverty measure between two populations and data errors are generated by the models under consideration. The problem is formulated as follows: there are two populations, A and B, characterized by distributions F and G, respectively. Moreover, we assume the existence of upper bounds λ_A and λ_B on the proportion of data errors. We are interested in comparing, in terms of some ASP measure, the two populations.

Define the difference between the poverty measures corresponding to distribution F and G as $D=\Pi(F;z)-\Pi(G;z)$. Proposition 1.1 can be used to obtain informative, although not necessarily sharp, outer bounds on D given λ_A and λ_B.[8][9]

Proposition 6 *Let it be known that* $p_A \leq \lambda_A < 1$ *and* $p_B \leq \lambda_B < 1$. *If* $\Pi(P; z)$ *belongs to the family of additively separable poverty measures and the poverty evaluation function is non-increasing in* y, *then identification regions for* $D(F_{11}, G_{11}; z)$ *and* $D(F_1; G_1; z)$ *are given by*

$$\mathbf{H}[D(F_{11}, G_{11}; z)] = [\Pi^l_{\lambda_A}(F; z) - \Pi^u_{\lambda_B}(G; z), \Pi^u_{\lambda_A}(F; z) - \Pi^l_{\lambda_B}(G; z)] \tag{25}$$

and

$$\mathbf{H}[D(F_1, G_1; z)] = [D_1^l, D_1^u] \tag{26}$$

where

$$D_1^u = (1 - \lambda_A)\Pi^u_{\lambda_A}(F; z) - (1 - \lambda_B)\Pi^l_{\lambda_B}(G; z) + \lambda_A \psi_1 - \lambda_B \psi_0$$
$$D_1^l = (1 - \lambda_A)\Pi^l_{\lambda_A}(F; z) - (1 - \lambda_B)\Pi^u_{\lambda_B}(G; z) + \lambda_A \psi_0 - \lambda_B \psi_1$$

7.1 Statistical Inference

Let $y_1, \ldots, y_n$ and $y_1, \ldots, y_m$ be two independent random samples drawn from F and G, respectively. We will construct confidence intervals for the identification region of the poverty difference $\Pi_A - \Pi_B$.

Define the confidence interval $CI_\gamma^{D_l, D_u}$ as follows

$$CI_\gamma^{[D_l, D_u]} = \left[\hat{\Pi}_l(F) - \hat{\Pi}_u(G) - z_{\frac{\gamma+1}{2}}\hat{\sigma}^*, \hat{\Pi}_u(F) - \hat{\Pi}_l(G) + z_{\frac{\gamma+1}{2}}\hat{\sigma}^{**}\right] \tag{27}$$

[8] In principle, it is not necessary to restrict both distributions to have same type of data errors. For instance, distribution A could be characterized by contaminated data while distribution B by corrupted data. The analysis and conclusions would not change by including that level of detail.

[9] As noticed by Manski (2003), outerbounds on differences between parameters that respect stochastic dominance are generally non-sharp. In the present case, for these to be sharp, there would have to exist two distributions of errors that jointly make $\Pi(F; z)$ and $\Pi(G; z)$ attain their sharp bounds.

where

$$\hat{\sigma}^{*} = \sqrt{\frac{\hat{\sigma}^{2}_{lF}}{n} + \frac{\hat{\sigma}^{2}_{uG}}{m}}$$
$$\hat{\sigma}^{**} = \sqrt{\frac{\hat{\sigma}^{2}_{uF}}{n} + \frac{\hat{\sigma}^{2}_{lG}}{m}}$$

Proposition 7 *Let* $\lambda_i < 1$, $i = A, B$ *be known,. Assume* $E_i(\pi(y; z)^2) < \infty$. *Let* $r_i(1-\lambda_i)$ *and* $r_i(\lambda_i)$ *be continuity points and let* $m, n \to \infty$ *such that* $\frac{m}{m+n} \to \epsilon \in (0, 1)$. *Let the poverty evaluation function,* $\pi(y; z)$, *be continuous at* $r_i(1-\lambda_i)$ *and* $r_i(\lambda_i)$. *Then*

$$\lim_{n,m\to\infty} \mathbb{P}([D_l, D_u] \subset CI_{\gamma}^{[D_l,D_u]}) \geq \gamma \tag{28}$$

PROOF: See Appendix.

8 Application: Evaluation of an Anti-Poverty Program with Missing Treatments

8.1 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. This 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.

Because of logistical and financial constraints, the program was introduced in several phases. The sequentiality of the program 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. However, as noticed by Buddelmeyer and Skoufias (2004), in the treatment localities 27% of the eligible population had not received any benefits by March 2000 due to some administrative error.

8.2 Poverty Treatment Effects

Let us introduce some basic notation that will be helpful for the rest of the section. There are two potential states of the world, (y_1, y_0), for each individual, where y_1 and y_0 are the outcomes that an individual would obtain if she were and she were not, respectively, a beneficiary of PROGRESA. Lets denote observed outcome by y and program participation by the indicator variable d, where $d = 1$ if the individual participates in the program, and $d = 0$ otherwise. The policymaker observes (y, d), but he cannot observe both states (y_1, y_0). Formally, the policymaker observes the random variable $y = dy_1 + (1 - d)y_0$.

We are interested in the poverty treatment effect (PTE) on the treated. This effect is given by:

$$\Delta = \Pi(F(y_1 \mid d = 1); z) - \Pi(F(y_0 \mid d = 1); z)$$

Where $F(y_1 \mid d = 1)$ is the distribution of the outcome of interest for the treated group, and $F(y_0 \mid d = 1)$ is its counterfactual. Randomization guarantees the identification of PTE since we have $F(y_0 \mid d = 1) = F(y_0 \mid d = 0)$.

As it was mentioned above, in the case of PROGRESA we have a problem of measurement error for the treatment group since a *proportion* of the households selected as beneficiaries had not received the cash transfer by the year 2000. Applying the model of section 3 to the current setting, let each individual in the treatment group be characterized by the tuple (y_{11}, y_{10}), where y_{11} and y_{10} are the outcomes that an individual randomized in the treatment group would obtain if she were and she were not, respectively, participating in PROGRESA. Instead of observing y_{11}, one observes a contaminated variable y_1 defined by

$$y_1 \equiv wy_{11} + (1 - w)y_{10} \tag{29}$$

From section 3 we know that $F(y_{11}) = F(y_1 \mid d = 1)$ cannot be point identified if $E(w) < 1$. However, it can be partially identified if we possess some information on the marginal probability of data errors $p = P(w = 0)$, in particular if there exists a non trivial upper bound on this probability.

If one assume that w is independent of y_{11}, which is equivalent to say that data from the treatment group is contaminated, then we can apply the results obtained in section 4 to find the identification region for the PTE:

$$\mathbf{H}[\Delta] = [\Pi^l_\lambda(F(y_{11}); z) - \Pi(F(y_0 \mid d = 1); z), \Pi^u_\lambda(F(y_{11}); z) - \Pi(F(y_0 \mid d = 1); z)]$$

where $\lambda < 1$ is as an upperbound on the probability of not receiving treatment when the unit of analysis has been randomized in the treatment group.

Under the assumptions of proposition 7, the following confidence interval, $CI^{[\Delta_l,\Delta_u]}_\gamma$, asymptotically covers the PTE with at least probability γ:

$$\left[\hat{\Pi}_l(F_{11}) - \hat{\Pi}(F_0) - z_{\frac{\gamma+1}{2}}\sqrt{\frac{\hat{\sigma}^2_{lF}}{n} + \frac{\hat{\sigma}^2_{F_0}}{m}}, \hat{\Pi}_u(F_{11}) - \hat{\Pi}(F_0) + z_{\frac{\gamma+1}{2}}\sqrt{\frac{\hat{\sigma}^2_{F_{11}}}{n} + \frac{\hat{\sigma}^2_{F_0}}{m}}\right]$$

Table 1.1 presents an application of the present analysis to the PROGRESA data set. Column 1 introduces a parameter measuring the severity of poverty for the FGT poverty measure described below. We use consumption as welfare indicator, and the poverty line z is set equal to the median consumption for the control group. We use an upper bound on the proportion of errors of 0.27, the proportion of households who had not received benefits from Progresa by 2000. Columns 2 and 3 presents treatment effects on poverty and 95% confidence intervals for this parameter, respectively, without taking into consideration the contamination problem, that is to say, assuming that the parameter is point identified. Finally, columns 4,5, and 6 introduce, respectively, upper and lower bounds on the PTE, and Bonferroni confidence intervals for the identification region.

Table 1: Identification regions and confidence intervals for treatments effects on poverty: PROGRESA 1999

	Δ	$CI^{\Delta}_{0.95}$	Δ_l	Δ_u	$CI^{[\Delta_l,\Delta_u]}_{0.95}$
$\alpha = 0$	-.068	[-.083,-.053]	-0.278	0.092	[-0.296,0.105]
α =1	-0.039	[-.045,-.033]	-0.148	0.009	[-0.153,0.015]
$\alpha = 2$	-0.021	[-.025,-.017]	-0.076	0.000	[-0.079,0.005]

8.3 Monotone Treatment Response, Data Contamination, and Missing Treatments

Monotonicity assumptions have been applied in other places to exploit their identifying power. 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 the situation where no prior information exists. In a missing treatments environment, Molinari (2005b) shows that one can extract information from the observations for which treatment data are missing using monotonicity assumptions.

Given the design of PROGRESA, one should expect that the outcome of interest (in our case consumption per capita) increases with program participation. More formally, we should expect that $y_{11} \geq y_{10}$. We have the following result

Proposition 8 *Suppose that $y_{11} \geq y_{10}$. Let it be known that $p \leq \lambda < 1$. Then sharp bounds for $\Pi(P_{11}; z)$ and $\Pi(P_1; z)$ are given by the identification region*

$$[\Pi(U_\lambda; z), \Pi(Q(y); z)]$$

PROOF: See Appendix.

Table 1.2 introduces the effect of the monotonicity assumption on the identification region for PTE. Clearly, considering the monotonic effect of Progresa on the treated population improves the inferential analysis of PTE by considerably shrinking the identification region.

Table 2: Identification regions under monotonicity assumptions: PROGRESA 1999

	Δ_l	Δ_u	$CI_{0.95}^{[\Delta_l,\Delta_u]}$
$\alpha = 0$	-0.278	-.068	[-0.296,-.053]
$\alpha = 1$	-0.148	-0.039	[-0.153,-.033]
$\alpha = 2$	-0.076	-0.021	[-0.079,-.017]

9 Application: Measurement of Rural Poverty in Mexico

The methodology developed in this paper is applied to the data obtained from the 2002 *Encuesta Nacional de Ingreso y Gasto de los Hogares* (ENIGH) held by INEGI (2002). This household income and expenditure survey is one of a series of surveys that are carried out under the same days of each year using identical sampling techniques.

The households are divided into zones of high and low population density. Low density population zones are those areas with fewer than 2500 inhabitants. It is common to identify these areas as rural ones. The rest of the zones (those with more than 2500 inhabitants) are identified as urban areas. The sample is representative for both urban and rural areas and at the national level. For the purposes of this study, we will just concentrate on the rural sub-sample which includes 6753 observations.

We have considered the extreme poverty line for rural areas constructed by INEGI-CEPAL for the 1992 ENIGH, following the methodology applied by the Ministry of Social Development in Mexico (2002) to inflate both the poverty line and all of the data into August 2000 prices. The rural poverty line is equal to 494.77 monthly 2002 pesos. In this paper we have used per capita current disposable income as indicator of economic welfare.[10] It is divided into monetary and non-monetary income. The monetary sources include wages and salaries, entrepreneurial rents, incomes from

[10]Due to lack of information, a final transformation of the original data was required: we will assume that each household member obtains the same proportion of total income as the others.

cooperatives, transfers, and other monetary sources. Non-monetary incomes include gifts, autoconsumption, imputed rents, and payments in kind.

The identification regions and the three different 95% confidence intervals for the class of FGT poverty measures are presented for both the contamination and the corruption models in Tables 1.3 and 1.4, respectively. We have no estimate of the frequency of data errors in the sample, so we present a sensitivity analysis using different values of λ. The first confidence interval corresponds to the point identified case ($\lambda = 0$). It is based on the point estimator $\pm$ 1.96 times its standard error. The second confidence interval is equal to the estimator of the lower bound minus 1.96, and the estimator of the upper bound plus 1.96 times their standard errors. The third confidence interval is the adjusted interval for the parameter $\overline{C}_N$. We found that there is almost no difference between the last two types of confidence intervals, that is to say, between the confidence interval covering the entire identification region and the one that provides the appropriate coverage for the poverty measure.

10 Conclusions

This paper has introduced the problems of data contamination and data corruption into the context of poverty measurement. When a proportion of the data is measured with error, a poverty measure cannot be point identified. However, we have shown that for the class of additively separable poverty measures it is possible to find identification regions under very mild assumptions. In particular, if there is an upper bound on the proportion of errors, we can obtain identification regions that take the form of closed intervals.

We consider the problem of statistical inference when a poverty measure is not point identified. Two type of confidence intervals are applied in the present study. For the first type, we have developed Bonferroni's confidence intervals that cover the

Table 3: Identification regions and confidence intervals for FGT poverty measures under contamination model: Rural Mexico, 2002

λ	$\Pi^L_{\alpha\lambda}$	$\Pi^U_{\alpha\lambda}$	$CI^{\Pi}_{0.95}$	$CI^{[\Pi_L,\Pi_U]}_{0.95}$	$\overline{CI}^{\Pi}_{0.95}$
			$\alpha = 0$		
0.00	0.287	0.287	[0.276, 0.298]		
0.01	0.282	0.289		[0.271, 0.300]	[0.272, 0.299]
0.02	0.275	0.292		[0.265, 0.304]	[0.266, 0.302]
0.03	0.268	0.294		[0.257, 0.306]	[0.259, 0.304]
0.05	0.252	0.299		[0.241, 0.311]	[0.243, 0.309]
0.07	0.234	0.304		[0.223, 0.316]	[0.225, 0.314]
0.10	0.209	0.312		[0.198, 0.325]	[0.200, 0.323]
			$\alpha = 1$		
0.00	0.093	0.093	[0.089, 0.098]		
0.01	0.088	0.094		[0.084, 0.099]	[0.085, 0.098]
0.02	0.083	0.095		[0.079, 0.100]	[0.080, 0.099]
0.03	0.077	0.096		[0.074, 0.101]	[0.074, 0.100]
0.05	0.066	0.097		[0.062, 0.103]	[0.063, 0.102]
0.07	0.055	0.099		[0.052, 0.106]	[0.053, 0.105]
0.10	0.042	0.101		[0.039, 0.109]	[0.040, 0.108]
			$\alpha = 2$		
0.00	0.042	0.042	[0.040, 0.045]		
0.01	0.038	0.043		[0.036, 0.046]	[0.036, 0.045]
0.02	0.034	0.043		[0.032, 0.047]	[0.033, 0.046]
0.03	0.031	0.043		[0.029, 0.048]	[0.029, 0.047]
0.05	0.024	0.044		[0.022, 0.049]	[0.022, 0.048]
0.07	0.018	0.045		[0.016, 0.050]	[0.017, 0.050]
0.10	0.011	0.046		[0.010, 0.053]	[0.011, 0.052]

entire identification region with some fixed probability. The second type applies and extends the results of Imbens and Manski (2004) by covering the true value of a poverty measure with at least some fixed probability. We also consider the problem of poverty comparisons, extending the methodology developed in the first part of the paper to a setting where two populations are compared in terms of poverty.

The results obtained in the paper are illustrated by means of two applications. The first application analyzes the effect of contaminated data on poverty treatment effects for an anti-poverty program in Mexico. The second application is a sensitivity analysis for the measurement of rural poverty in Mexico under different degrees of

Table 4: Identification regions and confidence intervals for FGT poverty measures under corruption model: Rural Mexico, 2002

λ	$\Pi^L_{\alpha\lambda}$	$\Pi^U_{\alpha\lambda}$	$CI^{\Pi}_{0.95}$	$CI^{[\Pi_L,\Pi_U]}_{0.95}$	$\overline{CI}^{\Pi}_{0.95}$
			$\alpha = 0$		
0.00	0.287	0.287	[0.276, 0.298]		
0.01	0.279	0.296		[0.268, 0.307]	[0.270, 0.306]
0.02	0.270	0.307		[0.259, 0.318]	[0.261, 0.316]
0.03	0.260	0.316		[0.250, 0.327]	[0.251, 0.325]
0.05	0.239	0.334		[0.229, 0.345]	[0.231, 0.344]
0.07	0.218	0.352		[0.208, 0.364]	[0.209, 0.362]
0.10	0.188	0.381		[0.179, 0.393]	[0.180, 0.391]
			$\alpha = 1$		
0.00	0.093	0.093	[0.089, 0.098]		
0.01	0.087	0.103		[0.083, 0.108]	[0.084, 0.107]
0.02	0.081	0.113		[0.077, 0.118]	[0.078, 0.117]
0.03	0.075	0.123		[0.071, 0.128]	[0.072, 0.127]
0.05	0.063	0.142		[0.059, 0.148]	[0.060, 0.147]
0.07	0.051	0.162		[0.048, 0.168]	[0.049, 0.167]
0.10	0.038	0.191		[0.035, 0.198]	[0.036, 0.197]
			$\alpha = 2$		
0.00	0.042	0.042	[0.040, 0.045]		
0.01	0.038	0.052		[0.036, 0.055]	[0.036, 0.055]
0.02	0.034	0.062		[0.032, 0.066]	[0.032, 0.065]
0.03	0.030	0.072		[0.028, 0.076]	[0.028, 0.075]
0.05	0.022	0.092		[0.021, 0.097]	[0.021, 0.096]
0.07	0.016	0.112		[0.015, 0.117]	[0.015, 0.116]
0.10	0.010	0.141		[0.009, 0.147]	[0.010, 0.146]

data contamination and data corruption. The two empirical applications show the importance of considering these types of data errors, when it is pertinent, to get a more accurate measurement of the phenomenon of poverty.

11 Appendix: Proofs

Proof of Proposition 1.1: We need to show that $\Pi(U_\lambda; z) \leq \Pi(P; z)$ and $\Pi(L_\lambda; z) \geq \Pi(P; z)$ for all $P \in P_\lambda$. Set $\psi(y; z) = -\pi(y; z)$, so $\psi(y; z)$ is a non-decreasing function. By lemma 1.1, it suffices to prove that U_λ stochastically dominates every member of P_λ and L_λ is stochastically dominated by every member of that set. The rest of the proof is identical to proposition 4 in Horowitz and Manski (1995)$\square$

Proof of Proposition 1.2: Take any probability distribution Q in $\mathcal{P}$. Clearly, the identification breakdown point for the head-count ratio is given by

$$\lambda^H = \min\{H_Q, 1 - H_Q\}$$

Since $\pi_j(y, z) = 0$ for all $y \geq z$ and $j \in \mathcal{D}$, we have that $\lambda_j^\psi = H_Q$ for all poverty measures in $\mathcal{D}$. Next, I claim that $\lambda_j^\phi \geq 1 - H_Q$. Assume, towards a contradiction, that $\lambda_j^\phi < 1 - H_Q$. Define $\lambda^* = \frac{\lambda_j^\phi + 1 - H_Q}{2}$ and let $\delta(c_j)$ be the Dirac measure at c_j. Clearly, we have

$$\Pi_j(L_{\lambda_j^\phi}; z) \leq \Pi_j(L_{\lambda^*}; z) \leq \Pi_j(L_{1-H_Q}; z) \leq \Pi_j(\delta(c_j); z) = c_j$$

A contradiction. Hence, $\{\lambda_j^\phi, \lambda_j^\psi\} \geq \{H_Q, 1 - H_Q\}$ for all $j \in \mathcal{D}$, and the result follows.$\square$

Lemma 2 *Let P_1 and P_2 be two probability measures on $(\mathbb{R}, \mathcal{B})$, with $\mathcal{B}$ the Borel sets of $\mathbb{R}$. Define the sets $\mathbf{A}_1$, $\mathbf{A}_2$, $\mathbf{A}_3$, where $\mathbb{R} = \mathbf{A}_1 \cup \mathbf{A}_2 \cup \mathbf{A}_3$, $\sup \mathbf{A}_2 \leq \inf \mathbf{A}_3$, $\mathbf{A}_1 \cap \mathbf{A}_i = \varnothing$, $i = 2, 3$, and*

$$\Lambda = \{(P_1, P_2) : P_1(A) = P_2(A), \forall A \in \mathcal{B} \cap \mathbf{A}_1; P_1(\mathbf{A}_3) = P_2(\mathbf{A}_2) = 0\}$$

Let $\delta(x) : \mathbb{R} \to \mathbb{R}$ be a measurable function and suppose there exists some $z \in \mathbf{A}_1 \cup \mathbf{A}_2$ with $\delta(x) = 0$ for all $x \geq z$, and $\delta(x) \geq 0$, otherwise. Then:

i) F_2 first order stochastically dominates F_1, where F_2 and F_1 are the distribution functions implied by probability measures P_2 and P_1, respectively.

ii) $E_{P_2}(\delta(x)) \leq E_{P_1}(\delta(x)),\ \forall (P_1, P_2) \in \Lambda$

PROOF:

i) Straightforward

ii) Because $\delta(x) = 0$ for all $x \in \mathbf{A}_3$, we have:

$$E_{P_1}(\delta(x)) = \int_{\mathbf{A}_1} \delta(x) dP_1 + \int_{\mathbf{A}_2} \delta(x) dP_1$$
$$E_{P_2}(\delta(x)) = \int_{\mathbf{A}_1} \delta(x) dP_2 + \int_{\mathbf{A}_2} \delta(x) dP_2$$

Therefore, $E_{P_2}(\delta(x)) \leq E_{P_1}(\delta(x))$ iff

$$\int_{\mathbf{A}_1} \delta(x) dP_2 + \int_{\mathbf{A}_2} \delta(x) dP_2 \leq \int_{\mathbf{A}_1} \delta(x) dP_1 + \int_{\mathbf{A}_2} \delta(x) dP_1$$
$$\Leftrightarrow$$
$$\int_{\mathbf{A}_2} \delta(x) dP_2 \leq \int_{\mathbf{A}_2} \delta(x) dP_1$$

Since $P_2(\mathbf{A}_2) = 0$ and $\delta(x) \geq 0$ the result follows.$\square$

Definition 4 *A class $\mathcal{G}$ of subsets of Ω is called a λ-system if*

i) $\Omega \in \mathcal{G}$

ii) If $G_1, G_2 \in \mathcal{G}$ and $G_1 \supseteq G_2$ then $D_1 \setminus D_2 \in \mathcal{G}$.

iii) If $\{G_n\}$ is an increasing sequence of sets in $\mathcal{G}$, the $\bigcup_{n=1}^{\infty} G_n \in \mathcal{G}$

Lemma 3 *(Sierpinski 1928) If $\mathcal{F}$ is stable under finite intersections, and if $\mathcal{G}$ is a λ-system with $\mathcal{G} \supseteq \mathcal{F}$, then $\mathcal{G} \supseteq \sigma(\mathcal{F})$*

Proof of Proposition 1.3: Define $\delta(y) = \pi_2(y; z) - \pi_1(y; z)$. We have

$$\begin{aligned}
m_2 &= \int \pi_2(y; z) dL_\lambda - \int \pi_2(y; z) dU_\lambda \\
&= \int \pi_1(y; z) dL_\lambda + \int \delta(y) dL_\lambda - \int \pi_1(y; z) dU_\lambda - \int \delta(y) dU_\lambda \\
&= m_1 + \int \delta(y) dL_\lambda - \int \delta(y) dU_\lambda
\end{aligned}$$

By construction, there exist sets $\mathbf{A}_1, \mathbf{A}_2$ and $\mathbf{A}_3$ in $\mathbb{R}$ such that $\mathbb{R} = A_1 \cup A_2 \cup A_3$, $\sup \mathbf{A}_2 \leq \inf \mathbf{A}_3$, and $P_{L_\lambda}(\mathbf{A}_3) = P_{U_\lambda}(\mathbf{A}_2) = 0$. Moreover, $\delta(y) \geq 0$ for all y. By lemma 1.2, it suffices to show that $(P_{L_\lambda}, P_{U_\lambda}) \in \Lambda$. We have four cases: $\mathbf{A}_1 = [\min\{r(\lambda), r(1-\lambda)\}, \max\{r(\lambda), r(1-\lambda)\}]$, $\mathbf{A}_1 = (\min\{r(\lambda), r(1-\lambda)\}, \max\{r(\lambda), r(1-\lambda)\}]$,$\mathbf{A}_1 = [\min\{r(\lambda), r(1-\lambda)\}, \max\{r(\lambda), r(1-\lambda)\})$, and $\mathbf{A}_1 = (\min\{r(\lambda), r(1-\lambda)\}, \max\{r(\lambda), r(1-\lambda)\})$. I will analyze just first case. A similar argument works for the other three. Define the sets $\mathbf{A}_2 = (-\infty, \min\{r(\lambda), r(1-\lambda)\})$ $\mathbf{A}_3 = (\max\{r(\lambda), r(1-\lambda)\}, \infty)$, and $A_1 = [\min\{r(\lambda), r(1-\lambda)\}, \max\{r(\lambda), r(1-\lambda)\}]$. By inspection, we have $P_{L_\lambda}(A_3) = P_{U_\lambda}(A_2) = 0$. Let $\mathcal{B}(\mathbf{A}_1)$ be the Borel sigma-field on $\mathbf{A}_1$. I will show that $P_{L_\lambda}(A) = P_{U_\lambda}(A)$ for all $A \in \mathcal{B}(\mathbf{A}_1)$ by applying a generating class argument. Write $\mathcal{E}$ for the class of all intervals $(\min\{r(\lambda), r(1-\lambda)\}, t]$, with $t \in \mathbf{A}_1$. The following series of claims proves this result:

Claim 1: $\sigma(\mathcal{E}) = \mathcal{B}(A_1)$

Let $\mathcal{O}$ stand for the class of all open subsets of $\mathbf{A}_1$, so $\mathcal{B}(\mathbf{A}_1) = \sigma(\mathcal{O})$. Each interval $(\min\{r(\lambda), r(1-\lambda)\}, t]$ in $\mathcal{E}$ has a representation $\bigcap_{n=1}^{\infty}(\min\{r(\lambda), r(1-\lambda)\}, t + \frac{1}{n})$. $\sigma(\mathcal{O})$ contains all open intervals, and it is stable under countable intersections. Hence, $\mathcal{E} \subset \mathcal{B}(\mathbf{A}_1)$. On the other hand, each open interval (a, t) on A_1 has a representation $(a, t) = \bigcup_{n=1}^{\infty}(\min\{r(\lambda), r(1-\lambda)\}, t - \frac{1}{n}] \bigcap (\min\{r(\lambda), r(1-\lambda)\}, , a]^c$, so $\mathcal{O} \subset \sigma(\mathcal{E})$ and thus $\sigma(\mathcal{E}) = \mathcal{B}(A_1)$.

Claim 2: $\mathcal{D} = \{A \in \mathcal{B}(A_1) : P_{U_\lambda}(A) = P_{L_\lambda}(A)\}$ is a λ-system

i) $\mathcal{A}_1 \in \mathcal{D}$ follows from the fact that $P_{U_\lambda}(\mathcal{A}_1) = P_{L_\lambda}(\mathcal{A}_1)\}$. ii) Let $A_1, A_2 \in \mathcal{D}$. By the properties of a probability measure, $P_i(A_1 \bigcap A_2^c) = P_i(A_1) + P_i(A_2^c) - P_i(A_1 \bigcap A_2^c)$, $i = 1, 2$. $P_1(A_1 \bigcap A_2^c) = P_2(A_1 \bigcap A_2^c)$ follows after some algebraic manipulations. Finally, we need to show that $\mathcal{D}$ is closed under increasing limits. Let $\{A_n\}$ be an increasing sequence of sets in $\mathcal{D}$ and $A = \bigcup_{n=1}^{\infty} A_n$. Define a sequence of indicator functions $\{\mathbf{1}_{A_n}\}$. Clearly, this is a positive and increasing sequence of functions. By the Monotone Convergence Theorem

$$\begin{aligned}\lim_{n\to\infty} E_{P_{U_\lambda}}(\mathbf{1}_{A_n}) &= E_{P_{U_\lambda}}(1_A)\\ &= E_{P_{L_\lambda}}(1_A)\\ &= \lim_{n\to\infty} E_{P_{U_\lambda}}(\mathbf{1}_{A_n})\end{aligned}$$

hence $P_{U_\lambda}(A) = P_{L_\lambda}(A)$.

Claim 3: $\mathcal{D} \supseteq \mathcal{E}$

By inspection, $P_{L_\lambda}((\min\{r(\lambda), r(1-\lambda)\}, t]) = P_{U_\lambda}((\min\{r(\lambda), r(1-\lambda)\}, t])$ for all $t \in \mathbf{A}_1$.

Since $\mathcal{E}$ is stable under finite intersections, by lemma 1.3 and claims 1, 2, and 3 we have $\mathcal{D} \supseteq \sigma(\mathcal{E}) = \mathcal{B}(\mathbf{A}_1)$. Hence $\mathcal{D} = \mathcal{B}(\mathbf{A}_1)$.□

Proof of Corollary 1.1: By Proposition 1.2 it suffices to show that $\pi_1(y;z) \geq \pi_2(y;z)$ for all $y \in (0,z)$. By continuity and monotonicity of $\pi_1(y;z)$ on $[0,z]$ there exists $\lambda \in (0,1)$ such that $\pi_1(y;z) = \lambda\pi_1(0;z) + (1-\lambda)\pi_1(z,z)$ for all $y \in (0,z)$. Therefore

$$\begin{aligned}f \circ \pi_1(y;z) &= f(\lambda\pi_1(0;z) + (1-\lambda)\pi_1(z,z))\\ &\leq \lambda f \circ \pi_1(0;z) + (1-\lambda) f \circ \pi_1(z;z)\\ &\leq \lambda\pi_1(0;z) + (1-\lambda)\pi_1(z;z)\\ &= \pi_1(y;z)\end{aligned}$$

Where I have made use of the convexity of f.□

Proof of Corollary 1.3: Condition iv) is equivalent to have $\pi_1 = f \circ \pi_2$ with $f' > 0$ and $f'' > 0$ (Pratt 1964). The result follows from corollary 1.1.□

Before proving the rest of Lemmas and Propositions, we introduce a number of preliminary results. Let $y_1, y_2, \ldots, y_n$ be *i.i.d.* random variables with distribution function $F(y)$, and let $y_{(1)}, y_{(2)}, \ldots, y_{(n)}$ denote the order statistics of the sample. Consider the trimmed mean given by

$$S_n = \frac{1}{[(\beta - \alpha)n]} \sum_{i=[\alpha n]+1}^{[\beta n]} y_{(i)} \tag{30}$$

where $0 \leq \alpha < \beta \leq 1$ are any fixed numbers and $[\cdot]$ represents the greatest integer function. Let $r(\alpha)$ and $r(\beta)$ be continuity points of $F(y)$. Further, define

$$G(y) = \begin{cases} 0 & \text{if } y < r(\alpha) \\ \frac{F(y)-\alpha}{\beta-\alpha} & \text{if } r(\alpha) \leq y < r(\beta) \\ 1 & \text{otherwise} \end{cases}$$

and set

$$\mu = \int_{-\infty}^{\infty} y dG(y) \tag{31}$$

$$\sigma^2 = \int_{-\infty}^{\infty} y^2 dG(y) - \mu^2 \tag{32}$$

Lemma 4 *(Stigler 1973) Assume $E(y^2) < \infty$, then*

$n^{\frac{1}{2}}(S_n - \mu) \xrightarrow{d} N(0, (1-\alpha)^{-2}((1-\alpha)\sigma^2 + (r(\alpha) - \mu)^2\alpha(1-\alpha)))$ *if* $\beta = 1$.

$n^{\frac{1}{2}}(S_n - \mu) \xrightarrow{d} N(0, (\beta)^{-2}((\beta)\sigma^2 + (r(\beta) - \mu)^2\beta(1-\beta)))$ *if* $\alpha = 0$.

Lemma 5 *(de Wet 1976) Assume $E(|y|^3) < \infty$, then*

$sup \left| \mathbb{P}\left(\sqrt{N}\frac{(S_n-\mu)}{\sigma} < x\right) - \Phi(x)\right| \longrightarrow 0$ *if* $\beta = 1$.

$sup \left| \mathbb{P}\left(\sqrt{N}\frac{(S_n-\mu)}{\sigma} < x\right) - \Phi(x)\right| \longrightarrow 0$ *if* $\alpha = 0$.

Proof of Proposition 1.4: Define the events

$$A_n = \left\{\Pi_l : \Pi_l \geq \hat{\Pi}_l - z_{\frac{\gamma+1}{2}} \frac{\hat{\sigma}_l}{\sqrt{n}}\right\}$$

$$B_n = \left\{\Pi_u : \Pi_u \leq \hat{\Pi}_u + z_{\frac{\gamma+1}{2}} \frac{\hat{\sigma}_u}{\sqrt{n}}\right\}$$

From the definition of the confidence interval, $CI_\gamma^{[P_L,P_U]}$, and Bonferroni's inequality

$$\mathbb{P}([\Pi_l,\Pi_u] \subset CI_\gamma^{[\Pi_l,\Pi_u]}) = \mathbb{P}(A_n \cap B_n) \geq \mathbb{P}(A_n) + \mathbb{P}(B_n) - 1$$

By lemma 1.4 $\frac{\sqrt{n}(\hat{\Pi}_i - \Pi_i)}{\hat{\sigma}_i} \xrightarrow{d} \mathbf{N}(0,1)$, $i = u, l$. Thus

$$\lim_{n\to\infty} Pr([\Pi_l,\Pi_u] \subset CI_\gamma^{[\Pi_l,P_u]}) \geq \gamma$$

and the result follows.□

Proof of Proposition 1.5: The result is a direct consequence of lemma 1.5, and Lemma 4 in Imbens and Manski (2004). □

Lemma 6 *If $X_n \xrightarrow{d} X = N(\mu_1,\sigma_1^2)$ and $Y_m \xrightarrow{d} Y = N(\mu_2,\sigma_2^2)$, and if X_n is independent of Y_m for all n and m, then $X_n + Y_m \xrightarrow{d} N(\mu_1+\mu_2, \sigma_1^2+\sigma_2^2)$.*

Proof: Let $Z_{n,m} = X_n + Y_m$. By independence of X_n and Y_m, its characteristic function can be written as

$$\varphi_{Z_{n,m}}(u_1,u_2) = \varphi_{X_n}(u_1)\varphi_{Y_m}(u_2)$$

By the Uniqueness Theorem we have

$$\begin{aligned}\lim_{n,m\to\infty} \varphi_{X_n}(u_1)\varphi_{Y_m}(u_2) &= \exp(iu_1\mu_1 - \frac{u_1^2\sigma_1^2}{2})\exp(iu_2\mu_2 - \frac{u_2^2\sigma_2^2}{2}) \\ &= \exp(\sum_{j=1}^{2} iu_j\mu_j - \frac{1}{2}\sum_{j=1}^{2} u_j^2\sigma_j^2)\end{aligned}$$

This expression corresponds to the characteristic function of the random vector $Z = (X,Y)$, where Z is Gaussian. Moreover, X and Y are independent since $Cov(X,Y) = 0$. The result follows. □

Proof of Proposition 1.7: Define the events

$$A_{n,m} = \left\{\Pi_{lA} - \Pi_{uB} : z_{\frac{\gamma+1}{2}}\sqrt{\tfrac{m}{n+m}\hat{\sigma}_{lA} + \tfrac{n}{n+m}\hat{\sigma}_{uB}} \geq Y_{n,m}\right\}$$
$$B_{n,m} = \left\{\Pi_{uA} - \Pi_{lB} : -z_{\frac{\gamma+1}{2}}\sqrt{\tfrac{m}{n+m}\hat{\sigma}_{lA} + \tfrac{n}{n+m}\hat{\sigma}_{uB}} \leq W_{n,m}\right\}$$

Where $Y_{n,m} = \sqrt{\frac{nm}{n+m}}(\hat{\Pi}_{lA} - \hat{\Pi}_{uB} - \Pi_{lA} + \Pi_{uB})$ and $W_{n,m} = \sqrt{\frac{nm}{n+m}}(\hat{\Pi}_{uA} - \hat{\Pi}_{lB} - \Pi_{uA} + \Pi_{lB})$. Notice that Lemma 1.4 implies

i) $\lim_{n,m\to\infty} \sqrt{\frac{nm}{n+m}}(\hat{\Pi}_{iA} - \Pi_{iA}) = \lim_{n,m\to\infty} \sqrt{\frac{m}{n+m}}\sqrt{n}(\hat{\Pi}_{iA} - \Pi_{iA}) \xrightarrow{d} N(0, \epsilon\sigma^2_{iA})$, $i = l, u$

ii) $\lim_{n,m\to\infty} \sqrt{\frac{nm}{n+m}}(\hat{\Pi}_{iB} - \Pi_{iB}) = \lim_{n,m\to\infty} \sqrt{\frac{n}{n+m}}\sqrt{m}(\hat{\Pi}_{iB} - \Pi_{iB}) \xrightarrow{d} N(0, (1-\epsilon)\sigma^2_{iB})$, $i = l, u$

By applying Lemmas 1.4 and 1.7 together it is easy to show that $Y_{n,m} \xrightarrow{d} \mathbf{N}(0, \epsilon\sigma^2_{lA} + (1-\epsilon)\sigma^2_{uB})$ and $W_{n,m} \xrightarrow{d} \mathbf{N}(0, \epsilon\sigma^2_{uA} + (1-\epsilon)\sigma^2_{lB})$. By Bonferroni's inequality we have:

$$\mathbb{P}([D_l, D_u] \subset CI_\gamma^{[D_l,D_u]}) \geq \mathbb{P}(A_{n,m}) + \mathbb{P}(B_{n,m}) - 1$$

Hence $\lim_{n,m\to\infty} \mathbb{P}([D_l, D_u] \subset CI_\gamma^{[D_l,D_u]}) \geq \gamma$ □

Proof of Proposition 1.8: From Proposition 1 in Horowitz and Manski (1995)

$$P_{11}(y_1) \in \mathcal{P}_{11}(\lambda) \equiv \mathcal{P} \cap \left\{ \frac{Q(y) - \lambda\phi_{00}}{1-\lambda} : \phi_{00} \in \mathcal{P} \right\}$$

For all $x \in \mathbb{R}$, define the indicator functions $\mathbf{1}(y_{11} \leq x)$ and $\mathbf{1}(y_{10} \leq x)$. By the monotonicity assumption

$$\mathbf{1}(y_1 \leq x) \leq \mathbf{1}(y \leq x)$$

Taking expectations at both sides of this inequality, we have that

$$P(y_1 \leq x) \leq Q(y \leq x)$$

for all $x \in \mathbb{R}$. This imposes a restriction on the set $\mathcal{P}_{11}(\lambda)$ since all of the distributions in this set must stochastically dominate the observed distribution $Q(y)$. Hence

$$Max \left\{ 0, \frac{Q(y \leq x) - \lambda\phi_{00}(y_0 \leq x)}{1-\lambda} \right\} \leq Q(y \leq x)$$

for all $x \in \mathbb{R}$. After some algebraic manipulations, we obtain that $Q(y \leq x) \leq$

$\phi_{00}(y_0 \leq x)$, which provides a restriction on the set of feasible distributions ϕ_{00}. Define the set of distribution functions stochastically dominated by $Q(y)$ by

$$\mathfrak{D} = \{\phi_{00} \in \mathcal{P} : \phi_{00}(y_0 \leq x) \geq Q(y \leq x), \forall x\}$$

one can characterize the identification region for the distribution $F(y_{11})$ under the monotonicity assumption as follows:

$$P(y_{11}) \in \mathcal{P}_{11}^{M}(\lambda) \equiv \mathcal{P} \cap \left\{ \frac{Q(y) - \lambda\phi_{00}}{1-\lambda} : \phi_{00} \in \mathfrak{D} \right\}$$

To prove the proposition, we just need to show that $Q(y) \in \mathcal{P}_{11}^{M}$ and that this distribution is stochastically dominated by all other distributions in $\mathcal{P}_{11}^{M}$ The first condition is trivially satisfied by defining $\phi_{00} = Q(y)$, and hence we have that $Q(y) \in \mathcal{P}_{11}^{M}(\lambda)$. Next, assume, towards a contradiction, that there exists some distribution in $\mathcal{P}_{11}^{M}(\lambda)$ that does not stochastically dominate $Q(y)$. Then, for some $x \in \mathbb{R}$ and some $\phi'_{00} \in \mathfrak{D}$, we have

$$Min \left\{ 1, \frac{Q(y \leq x) - \lambda\phi'_{00}(y_0 \leq x)}{1-\lambda} \right\} > Q(y \leq x)$$

From where $Q(y \leq x) > \phi'_{00}(y_0 \leq x)$, or $1 > Q(y \leq x) > 1 + \lambda(1 - \phi_{00})$, a contradiction since $\phi'_{00} \in \mathfrak{D}$ and ϕ_{00} is a probability measure.$\square$

References

Buddelmeyer, H. and E. Skoufias (2004). An evaluation of the performance of regression discountinuity design on progresa. World Bank Policy Research Working Paper, 3386.

Chakravarty, S. (1983). A new index of poverty. *Mathematical Social Sciences 6*, 761–766.

Chesher, A. and C. Schluter (2002). Welfare measurement and measurement error. *Review of Economic Studies 69*, 357–378.

Clark, S., R. Hemming, and D. Ulph (1981). On indices for the measurement of poverty. *Economic Journal 91*, 307–313.

Cowell, F. A. and M.-P. Victoria-Feser (1996). Poverty measurement with contamited data: A robust approach. *European Economic Review 40*, 1761–1771.

de Wet, T. (1976). Berry-esseen results for the trimmed mean. *S. Afr. Statist. J. 10*, 77–96.

Dominitz, J. and R. Sherman (2005). Identification and estimation of bounds on school performance measures: a nonparametric analysis of a mixture model with verification. Mimeo, California Institute of Technology.

Foster, J., J. Greer, and E. Thorbecke (1984). A class of decomposable poverty measures. *Econometrica 52*, 761–766.

Hagenaars, A. (1987). A class of poverty indices. *International Economic Review 28*, 583–607.

Hampel, F. (1974). The influence curve and its role in robust estimation. *Journal of the American Statistical Association 69*, 383–393.

Hampel, F., E. Ronchetti, P. Rousseeuw, and W. Stahel (1986). *Robust Statistics.* Wiley, New York.

Horowitz, J. E. and C. F. Manski (1995). Identification and robustness with contaminated and corrupted data. *Econometrica 63*, 281–302.

Huber, P. (1981). *Robust Statistics.* Wiley, NY.

Imbens, G. and C. Manski (2004). Confidence intervals for partially identified parameters. *Econometrica 72*, 1845–1857.

INEGI (2002). *Encuesta Nacional de Ingresos y Gastos de los Hogares de 2002*. Instituto Nacional de Estadistica, Geografia, e Informatica, Mexico.

Kakwani, N. (1980). On a class of poverty measures. *Econometrica 48*, 437–446.

Kakwani, N. (1993). Statistical inference in the measurement of poverty. *The Review of Economics and Statistics*, 632–639.

Kreider, B. and J. Pepper (2004). Inferring disability status from corrupt data. Mimeo.

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

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

Molinari, F. (2005a). Identification of probability distributions with misclassified data. Mimeo, Cornell University.

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

Pratt, J. (1964). Risk aversion in the small and the large. *Econometrica 32*, 122–136.

Ravallion, M. (1994). Poverty rankings using noisy data on living standards. *Economics Letters 45*, 481–485.

SEDESOL (2002). Nota tecnica para la medicion de la pobreza con base en los resultados de la encuesta nacional de ingresos y gastos de los hogares, 2002. Secretaria de Desarrollo Social, Mexico.

Sen, A. (1976). Poverty: An ordinal approach to measurement. *Econometrica 44*, 219–231.

Sierpinski, W. (1928). Un theoreme general sur les familles d'ensembles'. *Fundamenta Mathematicae 12*, 206–210.

Stigler, S. (1973). The asymptotic distribution of the trimmed mean. *The Annals of Statistics 1*, 472–477.

Szekely, M., N. Lustig, M. Cumpa, and J. Mejia (2000). Do we know how much poverty there is. Mimeo, Inter-American Development Bank.

Watts, H. (1968). An economic definition of poverty. In D. Moynihan (Ed.), *On Understanding Poverty*. New York: Basic Books.

OTHER A.E.M. WORKING PAPERS

WP No	Title	Fee (if applicable)	Author(s)
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
2005-20	Membership Based Organizing of Poor Women: Reflections After an Exposure and Dialogue Program with SEWA in Gujarat, India, January 2005		Editors: Chen, M., Jhabvala, R., Kanbur, R., Mirani, N., Osner, K. and C. Richards
2005-19	Oil Revenue, the Real Exchange Rate and Sectoral Distortion in Angola		Kyle, S.

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.