Partial Identification of Poverty Measures with Contaminated Data

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

April 14, 2004

Abstract

Much of the statistical analysis for poverty measurement regards the data employed to estimate poverty statistics as error-free observations. However, it is amply recognized that surveys responses are not perfectly reliable and that the quality of the data is often poor, especially for developing countries. Robust estimation addresses this problem by searching for poverty measures that are not highly sensitive to errors in the data. However, given the assumptions of robust estimation, the rationale for point estimation is not apparent. In the present study we tackle the problem by implementing a different strategy. Since a particular poverty measure is not point identified under the assumptions of robust estimation and some outcomes that are possible ex ante are ruled out ex post, we apply a fully non-parametric method to show that for the family of additively separable poverty measures it is possible to find identification regions under very mild assumptions. We investigate the sensitivity of the bounds of these identification regions to contamination for the class of P_α poverty measures, showing that there exists an "α-ordering" for the elasticities of these bounds with respect to the amount of contamination. We apply two conceptually different confidence intervals for partially identified poverty measures: the first type of confidence interval covers the entire identification region, while the other covers each element of the identification region with fixed probability. The methodology developed in the paper is applied to analyze rural poverty in Mexico.

JEL Classification: C14, I32.
Keywords: *Poverty Measurement, Bounds, Partial Identification, Contamination Model, Identification Regions, Confidence Intervals.*

*I am grateful to Francesca Molinari for helpful comments and suggestions, and Consejo Nacional de Ciencia y Tecnologia (CONACYT) and the Ford/MacArthur/Hewlett Program of Graduate Fellowships in the Social Sciences for financial support. All remaining errors are my own.

1 Introduction

Since the appearance of Sen's seminal paper [21], research on poverty measurement has focused on the theoretical properties of aggregate poverty measures. Much of the statistical analysis of poverty measurement regards the data employed to estimate a specific poverty measure as error-free observations, implicitly assuming that the real problem to be concerned about is sample size[1]. However, it is amply recognized that surveys responses are not perfectly reliable. Financial and technological constraints may affect the quality of the data, something that is particularly relevant for developing countries, making "truth" very difficult to measure. [1, 24, 22]

Measurement error has several dimensions for poverty estimation. For example, the poverty line is set for heterogenous groups of people without considering idiosyncratic differences in the cost of basic needs [18], arbitrary imputations are made when missing and zero outcomes appear in the sample, and the variable of interest is misreported by an important subset of survey respondents. [25]

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 [25] applied several techniques to adjust for misreporting. 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 set about this problem in a more structured way. One of these approaches, robust estimation, aims at developing point estimators that are not highly sensitive to errors in the data.[2]. The objective is to guard against worse-case scenarios that errors in the data could conceivably produce. In that

[1] There are some exceptions. For example, Chesher and Schulter [2] investigate the sensitivity of poverty measures to alternative amounts of measurement error.

[2] See Hampel et al [8] and Huber [10] for a comprehensive treatment of robust inference.

sense it takes an ex-ante perspective of the problem. Cowell and Victoria-Feser [3] apply this approach to poverty measurement by using the concept of the influence function: a statistical tool to assess the influence of an infinitesimal amount of contamination upon the value of a statistic [7]. They found that poverty measures that take as their primitive concept poverty gaps rather incomes of the poor are in general robust under this criterion. In particular, they proved that the class P_α of poverty measures developed by Foster, Greer and Thorbecke [6] is robust under data contamination.

In the present study we tackle the problem by implementing a different strategy. Since the population parameters of interest are not point identified under the assumptions of robust estimation and some outcomes that are possible ex ante are ruled out ex post, we follow Horowitz and Manski [9] and apply a partial identification approach for poverty measurement[3]. 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 poverty measures belonging to the additively separable class. It is shown that, by using some stochastic dominance properties, we can find upper and lower bounds for poverty measures within that class. In section 5, we analyze the sensitivity of the bounds to contamination for the class of P_α poverty measures, showing that there exists an "α-ordering" for the elasticities of these bounds with respect to contamination. Section 6 applies two conceptually different kinds of confidence intervals for partially identified poverty measures. Section 7 provides an empirical illustration by applying the methodology to the measurement of rural poverty in Mexico. Section 8 concludes. Most of the mathematical details are in the Appendix.

[3]Examples of applications of this approach in other settings are Molinari [17] and Dominitz and Sherman [5]. See Manski [16] for an overview of this literature

2 Poverty Measurement: Conceptual Framework

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

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

Definition 1 *A Poverty Index is a functional* $\Pi(P; z) : \mathcal{P} \times \Re_{++} \to \Re$ *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* class[4], which is defined as follows:

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

Where $\pi(y; z) : \Re_{++} \times \Re \to \Re$, is the poverty evaluation function for an individual, indicating 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 [21], 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 second axiom proposed by Sen says that, everything else equal, a reduction in the income of a poor individual must increase the poverty measure; formally:

MONOTONICITY AXIOM: If $y_1 < y_2$, then $\pi(z; y_1) > \pi(z; y_2)$

Sen's third axiom emphasizes the positive effect of a regressive transfer on the poverty measure:

TRANSFER AXIOM: *Given other things, a pure transfer of income from a poor individual to any other individual that is richer must increase the poverty measure.*

[4]Members of this class of poverty measures are the FGT, the Watts, and the Clark, Hemming and Ulph poverty measures. See Seidl [20] for a survey of poverty measures.

Kakwani [13] has proposed a 4th property that emphasizes transfers taking place down in the distribution, other things being equal; formally:

TRANSFER SENSITIVITY AXIOM: *If a transfer $t > 0$ of income takes place from a poor individual with income y to a poor individual with income $y + \delta$ ($\delta > 0$), then the magnitude of the increase in poverty must be smaller for larger y_i.*

3 Statement of the Problem

Let each member j of population J be characterized by the pair of outcomes (y_1^j, y_0^j) in the space $\Re \times \Re$ where y_1^j denote the "true" equivalent income (or expenditure) for a given poverty line z. Let the random variable $(y_1, y_0) : J \longrightarrow \Re \times \Re$ 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:

$$y \equiv wy_1 + (1 - w)y_0 \tag{2}$$

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 $i = 0, 1$ and $j = 0, 1$. 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:

$$P_1 = (1 - p)P_{11} + pP_{10} \tag{3}$$

$$Q(y) = (1 - p)P_{11} + pP_{00} \tag{4}$$

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 identification analysis $Q(y)$ is held 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:

$$P_1 = P_{11} \tag{5}$$

This particular model of data errors is known as "contaminated data" or "contaminated sampling" model. [10] In the other case, (5) does not hold and it is only assumed that there exists a non-trivial upper bound on the error probability. Horowitz and Manski [9] refer to this case as "corrupted sampling".

Define the sets

$$\mathcal{P}_1(p) \equiv \mathcal{P} \cap \{(1-p)\phi_{11} + p\phi_{10} : (\phi_{11}, \phi_{10}) \in \mathcal{P}_{11}(p) \times \mathcal{P}\} \tag{6}$$

$$\mathcal{P}_{11}(p) \equiv \mathcal{P} \cap \left\{ \frac{Q - p\phi_{00}}{1-p} : \phi_{00} \in \mathcal{P} \right\} \tag{7}$$

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. [9]

4 Partial Identification of Poverty Measures

Suppose now that a proportion $p < 1$ of the data is erroneous. Furthermore, assume there exists a non-trivial upper bound, λ, for p, so $p \leq \lambda < 1$.[5] From the analysis above, we know that the distribution of interest P_1 is not identified.

Even though P_1 is not 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 set of values in $\Re$ of a given poverty measure. So the natural question is if there is a way to bound such values. As we will see below, it is possible to do it for the class of additively separable poverty indices for which the poverty evaluation function is decreasing by ordering the distributions in $\mathcal{P}_\lambda$ according to a stochastic dominance criterion. Such criterion is defined as follows:

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

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

for all $x \in \Re$.

There is a well-known equivalent result for FOC that will be helpful to obtain some of the results in this study:

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

$$\int \varphi(x) dF(x) \geq \int \varphi(x) dG(x) \tag{8}$$

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

Following Horowitz and Manski [9] we can construct identification regions for ASP measures.

[5] 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 [15] for an application of a validation model.

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

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

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

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:*

$$\mathbf{H}[\Pi(P_{11}; z)] = [\Pi_l(U_\lambda; z), \Pi_u(L_\lambda; z)] \tag{9}$$

and

$$\mathbf{H}[\Pi(P_1; z)] = [(1-\lambda)\Pi_l(U_\lambda; z) + \lambda\psi_0, (1-\lambda)\Pi_u(L_\lambda; z) + \lambda\psi_1] \tag{10}$$

where ψ_0 *and* ψ_1 *are the lower and upper bounds of the poverty evaluation function respectively.*

Proof: See appendix.

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 Sensitivity of the Bounds to Contamination: An α-Ordering

Recent years have witnessed an emphasis in the axiomatic approach for poverty measurement. Following Sen [21], there has been a widely use of distributive-sensitive poverty measures. This new trend is epitomized by the class P_α of poverty measures developed by Foster et al [6], which is not only a member of the class of ASP poverty measures, but also one of the most widely-used in applied work.

Define $\Gamma = \{F(y) : F(y) = P((-\infty, y]); P((-\infty, y]) = 0, \forall y < 0\}$, $\Gamma \in \mathcal{P}$, i.e. the support of y is on $\Re_+$. The P_α measure is given by

$$P_\alpha(F; z) = \int 1(y < z) \left(\frac{z-y}{z}\right)^{\alpha-1} dP \tag{11}$$

Where $\alpha \geq 1$ can be viewed as a measure of poverty aversion: The larger α, the greater the relative importance of the poorest individuals. Since P_α belongs to the class of ASP measures and its evaluation function is non-increasing in y, we can find its identification region in presence of contamination by Proposition 1. Define $\Phi(y; z) = 1(y < z)\left(\frac{z-y}{z}\right)^{\alpha-1}$, $P^L_{\alpha\lambda} = \int \Phi(y; z) dU_\lambda$, and $P^U_{\alpha\lambda} = \int \Phi(y; z) dU_\lambda$. From Proposition 1, the identification region for P_α when the data is contaminated is given by:

$$\mathbf{H}[P_\alpha] = \left[P^L_{\alpha\lambda}, P^U_{\alpha\lambda}\right] \tag{12}$$

It is relevant to investigate the effects of contaminated data on the bounds of the identification region when the FGT poverty measure is more distributive-sensitive: ie larger values of the parameter α. A natural conjecture would be that the bounds are more sensitive to contamination, the larger the parameter α is. An approach to verify this conjecture is to perform a relative comparison of the influence that infinitesimal changes in contamination, represented by the parameter λ, have on the bounds for different values of α. To accomplish this task, I will make use of the concept of elasticity.

Definition 4 *For the family P_α of poverty measures, the elasticities of the lower and upper bounds with respect to λ are defined, respectively, as*

$$\xi^L_{\alpha\lambda} = \left|\frac{\lambda \frac{dP^L_{\alpha\lambda}}{d\lambda}}{P^L_{\alpha\lambda}}\right| \tag{13}$$

$$\xi^U_{\alpha\lambda} = \left|\frac{\lambda \frac{dP^U_{\alpha\lambda}}{d\lambda}}{P^U_{\alpha\lambda}}\right| \tag{14}$$

Assume $F(x)$ is differentiable at $r(1-\lambda)$ and $r(\lambda)$ with $F'(r(1-\lambda)) \neq 0$ and $F'(r(\lambda)) \neq 0$.

Proposition 2 $\xi^U_{\alpha\lambda} \geq \xi^U_{\beta\lambda}$, *and* $\xi^L_{\alpha\lambda} \geq \xi^L_{\beta\lambda}$ *whenever* $\alpha > \beta \geq 1$

Proof: See appendix

Proposition 2 provides a very important insight: there is a positive relationship between the sensitivity of a poverty measure to the way income is distributed and the effects of changes in contamination on the bounds. This is particularly relevant if we think of the distributive considerations of the axiomatic approach. For example, it can be shown that for $\alpha > 1$, P_α satisfies the Monotonicity Axiom, the Transfer Axiom for $\alpha > 2$, and the Transfer Sensitivity axiom for $\alpha > 3$. Therefore, if the "precision" of a poverty estimator is measured by the "reaction" of the bounds to small changes in contamination, its usefulness might be limited even if it possess some desirable properties.[6]

6 Confidence intervals for Partially Identified Poverty Measures

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

In the point identified case, a consistent estimator of the class of ASP measures is given by

$$\hat{\Pi} = \frac{1}{N}\sum_{i=1}^{N} 1(y_i < z)\pi(y_i; z) \tag{15}$$

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

$$CI^{\Pi}_{\gamma} = \left[\hat{\Pi} - z_{\frac{\hat{\gamma}+1}{2}}\frac{\sigma}{\sqrt{N}}, \hat{\Pi} + z_{\frac{\gamma+1}{2}}\frac{\hat{\sigma}}{\sqrt{N}}\right] \tag{16}$$

[6] Notice the difference between this problem and the construction of confidence intervals. For example, under the assumption of no data-errors, Kakwani [14] argues that "if the sample size is large enough, one can always estimate any poverty measure with precision". In identification analysis the issue of precision does not disappear even when the sample size is *large enough*.

where z_τ is the τ quantile of the standard normal distribution.[7]

I will apply two conceptually different methodologies to estimate confidence intervals when data is contaminated. The first methodology considers symmetric confidence intervals for the entire identification region $\mathbf{H}[\Pi(P_1; z)]$. The second type of confidence interval, developed by Imbens and Manski [11], rather than cover the entire identification region with fixed probability γ, asymptotically covers the true value of the parameter with this probability. Besides, this type of confidence interval ensures that its exact coverage probability does converge uniformly to its nominal values. By doing so, one is able to avoid the problem of having wider confidence intervals when the parameter is point identified that when is set-identified.

For the first class of confidence intervals, I will make use of a result on L-statistics due to Stigler [23], who explores the asymptotic behavior of trimmed means. Define the confidence interval $CI_\gamma^{[\Pi_L,\Pi_U]}$ as

$$CI_\gamma^{[\Pi_L,\Pi_U]} = \left[\hat{\Pi}_L - z_{\frac{\gamma+1}{2}}\frac{\hat{\sigma}_L}{\sqrt{n}}, \hat{\Pi}_U + z_{\frac{\gamma+1}{2}}\frac{\hat{\sigma}_U}{\sqrt{n}}\right] \tag{17}$$

Where $\hat{\sigma}_L^2$ and $\hat{\sigma}_U^2$ are, respectively, consistent estimators for

$$\sigma_L^2 = \frac{Var_{U_\lambda}(\pi(y;z)) + (\pi(r(1-\lambda)) - \Pi_L)\lambda}{1-\lambda} \tag{18}$$

$$\sigma_U^2 = \frac{Var_{L_\lambda}(\pi(y;z)) + (\pi(r(\lambda)) - \Pi_U)\lambda}{1-\lambda} \tag{19}$$

Proposition 3 *Let $\lambda < 1$ be known. Assume $E(\pi(y;z)^2) < \infty$. Let $r(1-\lambda)$ and $r(\lambda)$ be continuity points of $F(x)$. Let the poverty evaluation function, $\pi(y;z)$, be a non-increasing function that is continuous at $r(1-\lambda)$ and $r(\lambda)$. Then*

$$\lim_{n\to\infty} Pr([P_L, P_U] \subset CI^{[\Pi_L,\Pi_U]}\gamma) \geq \gamma \tag{20}$$

Proof: See appendix.

[7] Kakwani [14] describes this methodology for ASP indices

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 assumptions:

Assumption 1 *$F(y) \in \mathcal{F}$, where $\mathcal{F}$ is the set of distribution functions for which $E(\mid \pi(y;z) \mid^3) < \infty$ and $\underline{\sigma}^2 \leq \sigma_l^2, \sigma_u^2$ for some positive and finite $\underline{\sigma}^2$*

Assumption 2 $\Pi_u - \Pi_l \leq \overline{\Delta} < \infty$

Assumption 3 *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 $P \in \mathcal{P}$.*

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

$$\overline{CI}_{\gamma}^{\Pi} = \left[\widehat{\Pi}_l - \frac{\overline{C}_N \hat{\sigma}_l}{\sqrt{N}}, \widehat{\Pi}_u + \frac{\overline{C}_N \hat{\sigma}_u}{\sqrt{N}}\right] \tag{21}$$

where $\overline{C}_N$ satisfies

$$\Phi\left(\overline{C}_N + \sqrt{N}\frac{\hat{\Delta}}{max(\hat{\sigma}_l, \hat{\sigma}_u)}\right) - \Phi\left(-\overline{C}_N\right) = \gamma \tag{22}$$

Proposition 4 *Suppose assumptions 1,2 and 3 hold. Then*

$$lim_{N\rightarrow\infty} inf_{P\in\mathcal{P}} Pr\left(\Pi \in \overline{CI}_{\gamma}^{\Pi}\right) \geq \gamma \tag{23}$$

Proof: See appendix.

7 An Application to 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 [12]. 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 SEDESOL [19] 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[8] . It is divided into monetary and non-monetary income. The monetary sources include wages and salaries, entrepreneurial rents, incomes from 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 Figures 1 and 2 respectively[9]. The contamination model applies if the occurrence of events that produces data errors is statistically independent of y_1, the outcome of interest. The corruption model applies if the occurrence of those events is not statistically independent of y_1. 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 $\overline{C}_N$.

For this particular data set, 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 parameter of interest.

[8]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.

[9]We have no estimate of the frequency of data errors in the sample, so we present a sensitivity analysis using different values of λ.

It is clear from the empirical exercise that only considering random sampling errors without paying attention to the effects of measurement errors on poverty estimation is very likely to produce considerable bias in our inferences. For instance, it is remarkable how the lower bound is much more sensitive than the upper bound to changes in λ, so the event defined as "poverty is overestimated" cannot be discarded even if the sample size is "big enough", that is to say, the exact knowledge of a poverty measure cannot be inferred from any finite number of observations when the data is contaminated.

Figure 1: Identification regions and confidence intervals for FGT poverty measures under contamination model: Rural Mexico, 2002

λ	$P^L_{\alpha\lambda}$	$P^U_{\alpha\lambda}$	$CI^{\Pi}_{0.95}$	$CI^{[\Pi_L,\Pi_U]}_{0.95}$	$\overline{CI}^{\Pi}_{0.95}$
			$\alpha = 0$		
0	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	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	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]

Figure 2: Identification regions and confidence intervals for FGT poverty measures under corruption model: Rural Mexico, 2002

λ	$P^L_{\alpha\lambda}$	$P^U_{\alpha\lambda}$	$CI^{\Pi}_{0.95}$	$CI^{[\Pi_L,\Pi_U]}_{0.95}$	$\overline{CI}^{\Pi}_{0.95}$
			$\alpha = 0$		
0	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	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	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]

8 Conclusions

In the last decade a growing body of research has studied inference in settings where parameters of interest are not point identified. The main contribution of this paper has been to bring about this literature in the context of poverty measurement.

When the observed data is corrupted or contaminated and without making parametric assumptions on the distribution from which the data are drawn, a particular poverty measure is not point identified. By applying the work on contaminated and corrupted samples developed by Horowitz and Manski [9], and using some properties common to an important subset of poverty measures, we have been able to identify bounds for the class of additively separable poverty indexes. Moreover, we have shown that for the class of P_α poverty measures the bounds of its identification regions are more sensitive to changes in contamination, the larger the parameter α is (ie the more distributive-sensitive the poverty measure is).

We have extended two different confidence intervals to the setting of partially identified poverty measures. The first type of confidence interval provides coverage for the entire identification region, while the second one asymptotically covers the true value of the the poverty measure with fixed probability. We have illustrated the methodology developed in the paper with an application to rural poverty in Mexico. It is clear from both the theoretical and the empirical analysis that only considering random sampling errors without paying attention to the effects of measurement errors on poverty estimation is very likely to produce considerable bias in our inferences.

In future work, we plan to address questions about the identifying power of validation and covariate data, and monotonicity restrictions among other factors.

9 Appendix

Proof of Proposition 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, it is enough 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 [9]□

Before proving Proposition 2 we establish a preliminary result:

Lemma 2 *Let* $\Pi(F; z)$ *belong to the ASP class. If the monotonicity axiom holds, then the upper bound (lower bound),*$\Pi_u(L_\lambda; z)$ $(\Pi_l(U_\lambda; z))$*, is increasing (decreasing) with respect to* λ*. Formally*

$$\frac{d\Pi_u}{d\lambda} \geq 0$$

$$\frac{d\Pi_l}{d\lambda} \leq 0$$

Proof:

I will just prove the result for the upper bound since the proof for the lower bound is similar. From Proposition 1, the upper bound on $\mathbf{H}[\Pi(P_1; z)]$ is given by:

$$\Pi_u = \frac{1}{1-\lambda}\int_{-\infty}^{r(1-\lambda)} \pi(y; z)dQ$$

By Leibnitz's Rule

$$\frac{d}{d\lambda}\int_{-\infty}^{r(1-\lambda)} \pi(y; z)dQ = \pi(r(1-\lambda); z)$$

Whence

$$\frac{d}{d\lambda}\Pi_u = \frac{1}{1-\lambda}\left[\Pi u - \pi(r(1-\lambda); z)\right]$$

Define a Dirac measure at $r(1-\lambda)$

$$\delta(y) = \begin{cases} 0 & \text{if } y < r(1-\lambda) \\ 1 & \text{otherwise} \end{cases}$$

Hence

$$\pi(r(1-\lambda); z) = \int \pi(y; z)d\delta(y)$$

Notice that, for any $\lambda < 1$, $\delta(y)$ stochastically dominates all of the distribution functions defined by:

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

Since $\pi(y; z)$ is non-increasing on y, the result follows from lemma 1. □

Proof of Proposition 2: By Lemma 2, the derivative of $\int \Phi(y; z)dL_\lambda$ with respect to λ is positive for any $\alpha \geq 1$, the elasticities $\xi^U_{\alpha\lambda}$ and $\xi^U_{\beta\lambda}$ can be re-written as follows

$$\xi^U_{\alpha\lambda} = \frac{\lambda}{1-\lambda}\left[1 - \mathbf{1}(r(1-\lambda) < z)\frac{z^{1-\alpha}(z - r(1-\lambda))^{\alpha-1}}{P^U_{\alpha\lambda}}\right]$$

$$\xi^U_{\beta\lambda} = \frac{\lambda}{1-\lambda}\left[1 - \mathbf{1}(r(1-\lambda) < z)\frac{z^{1-\beta}(z - r(1-\lambda))^{\beta-1}}{P^U_{\beta\lambda}}\right]$$

Suppose (towards a contradiction) that $\xi^U_{\beta\lambda} > \xi^U_{\alpha\lambda}$. Equivalently

$$\mathbf{1}(r(1-\lambda) < z) \int_0^{r(1-\lambda)} \mathbf{1(y<z)} \left(\frac{z-y}{z-r(1-\lambda)} \right)^{\beta-1} dQ(y) > \mathbf{1}(r(1-\lambda) < z) \int_0^{r(1-\lambda)} \mathbf{1(y<z)} \left(\frac{z-y}{z-r(1-\lambda)} \right)^{\alpha-}$$

If $r(1-\lambda) \geq z$ this is a contradiction. Suppose $r(1-\lambda) < z$. Whence there exists $y^* \leq r(1-\lambda)$ such that:

$$\left(\frac{z-y^*}{z-r(1-\lambda)} \right)^{\beta-1} > \left(\frac{z-y^*}{z-r(1-\lambda)} \right)^{\alpha-1}$$

After some algebraic manipulations we get

$$r(1-\lambda) < y^*$$

which is a contradiction. The proof is analogous for the lower bound. □

Before proving Propositions 3 and 4, we present a number of preliminary results. Let $Y_1, Y_2, \ldots, Y_n$ be *i.i.d.* random variables with distribution function $F(y)$. Let $Y_{(1)}, Y_{(2)}, \ldots, Y_{(n)}$ denote the order statistics of the sample. Consider the trimmed mean given by

$$S_n = \frac{1}{(\beta-\alpha)n} \sum_{i=\alpha n+1}^{\beta n} Y_{(i)} \tag{24}$$

where $0 \leq \alpha < \beta \leq 1$. Let $F(r(\alpha)) = \alpha$ and $F(r(\beta)) = \beta$. 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

$$\mu = \int_{-\infty}^{\infty} y dG(y) \tag{25}$$

$$\sigma^2 = \int_{-\infty}^{\infty} y^2 dG(y) - \mu^2 \tag{26}$$

Lemma 3 *(Stigler [23]) Let $Y_1, Y_2, \ldots, Y_n$ be i.i.d. random variables with distribution function $F(y)$. then*

$$n^{\frac{1}{2}}(S_n - \mu) \xrightarrow{d} N(0, (1-\alpha)^{-2}((1-\alpha)\sigma^2 + (r(\alpha)-\mu)^2\alpha(1-\alpha))) \quad \text{if } \beta = 1 \text{ and } \int_0^{\infty} y^2 dF(y) < \infty$$
$$n^{\frac{1}{2}}(S_n - \mu) \xrightarrow{d} N(0, (\beta)^{-2}((\beta)\sigma^2 + (r(\beta)-\mu)^2\beta(1-\beta))) \quad \text{if } \alpha = 0 \text{ and } \int_{-\infty}^{0} y^2 dF(y) < \infty$$

Lemma 4 *(Berry-Esseen for trimmed means, de Wet [4]) Let $Y_1, Y_2, \ldots, Y_n$ be i.i.d. random variables with distribution function $F(y)$. Then*

$$sup\left|Pr\left(\sqrt{N}\frac{(S_n-\mu)}{\sigma} < x\right) - \Phi(x)\right| \longrightarrow 0 \quad \text{if } \beta = 1 \text{ and } \int_{r(\alpha)}^{\infty} |y|^3 dF(y) < \infty$$
$$sup\left|Pr\left(\sqrt{N}\frac{(S_n-\mu)}{\sigma} < x\right) - \Phi(x)\right| \longrightarrow 0 \quad \text{if } \alpha = 0 \text{ and } \int_{-\infty}^{r(\beta)} |y|^3 dF(y) < \infty$$

For Lemma 5 define $\Delta = \theta_u - \theta_l$ and let $\hat{\theta}_u$ and $\hat{\theta}_l$ and $\hat{\Delta} = \hat{\theta}_u - \hat{\theta}_l$ be estimators for θ_l, θ_u and Δ and consider the following set of assumptions:

i)There are estimators for the lower and upper bound $\widehat{\theta}_l$ and $\widehat{\theta}_u$ that satisfy: $\sqrt{N}(\hat{\theta}_l - \theta_l) \xrightarrow{d} \mathcal{N}(0, \sigma_l^2)$, and $\sqrt{N}(\hat{\theta}_u - \theta_u) \xrightarrow{d} \mathcal{N}(0, \sigma_u^2)$, uniformly in $P \in \mathcal{P}$ and there are estimators for σ_l^2 and σ_u^2 that converge to the true values uniformly in $P \in \mathcal{P}$.

ii) For all $P \in \mathcal{P}$, $\underline{\sigma}^2 \leq \sigma_l^2, \sigma_u^2 \leq \overline{\sigma}^2$ for some positive and finite $\underline{\sigma}^2$ and $\overline{\sigma}^2$, $\theta_u - \theta_l \leq \overline{\Delta} < \infty$.

iii)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 $P \in \mathcal{P}$.

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

$$\overline{CI}_{\gamma}^{\theta} = [\widehat{\theta}_l - \frac{\overline{C}_N \sigma_l}{\sqrt{N}}, \widehat{\theta}_u + \frac{\overline{C}_N \sigma_u}{\sqrt{N}}] \tag{27}$$

where $\overline{C}_N$ satisfies

$$\Phi(\overline{C}_N + \sqrt{N}\frac{\hat{\Delta}}{max(\hat{\sigma}_l, \hat{\sigma}_u)}) - \Phi(-\overline{C}_N) = \gamma \tag{28}$$

Lemma 5 *(Imbens and Manski, 2004) Suppose assumptions i), ii), and iii) hold. Then*

$$lim_{N\to\infty} inf_{P\in\mathcal{P}} Pr\left(\theta \in \overline{CI}_{\gamma}^{\theta}\right) \geq \gamma \tag{29}$$

Proof of Proposition 3:

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, $CI_\gamma^{[P_L, P_U]}$

$$Pr([\Pi_l, \Pi_u] \subset CI_\gamma^{[\Pi_l, \Pi_u]}) = Pr(A \cap B)$$

By Bonferroni's inequality:

$$Pr(A_n \cap B_n) \geq Pr(A_n) + Pr(B_n) - 1 \tag{30}$$

By lemma 3

$$\frac{\sqrt{n}(\hat{\Pi}_i - \Pi_i)}{\hat{\sigma}_i} \xrightarrow{d} \mathcal{N}(0, 1)$$

Therefore:

$$Pr([\Pi_l, \Pi_u] \subset CI^{[\Pi_l, P_u]}\gamma) \geq 2(\frac{\gamma + 1}{2}) - 1 = \gamma$$

asymptotically. □

Proof of Proposition 4:

The result is a direct consequence of lemmas 4 and 5. □

References

[1] J. Bound, C. Brown, and Nancy Mathiowetz. Measurement error in survey data. In J. Heckman and E. Leamer, editors, *Handbook of Econometrics, vol. 5.* Elsevier Science, 2001.

[2] Andrew Chesher and Christian Schluter. Welfare measurement and measurement error. *Review of Economic Studies*, 69:357–378, 2002.

[3] Frank A. Cowell and Maria-Pia Victoria-Feser. Poverty measurement with contamited data: A robust approach. *European Economic Review*, 40:1761–1771, 1996.

[4] T. de Wet. Berry-esseen results for the trimmed mean. *S. Afr. Statist. J.*, 10:77–96, 1976.

[5] J. Dominitz and R. Sherman. Nonparametric analysis of mixture models with verification. Mimeo, California Institute of Technology, 2003.

[6] James Foster, Joel Greer, and Erick Thorbecke. A class of decomposable poverty measures. *Econometrica*, 52:761–766, 1984.

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

[8] F. Hampel, E. Ronchetti, P. Rousseeuw, and W. Stahel. *Robust Statistics.* Wiley, New York, 1986.

[9] Joel E. Horowitz and Charles F. Manski. Identification and robustness with contaminated and corrupted data. *Econometrica*, 63:281–302, 1995.

[10] P. Huber. *Robust Statistics.* Wiley, NY, 1981.

[11] G. Imbens and C. Manski (forthcoming). Confidence intervals for partially identified parameters. *Econometrica.*

[12] INEGI. *Encuesta Nacional de Ingresos y Gastos de los Hogares de 2002.* Instituto Nacional de Estadistica, Geografia, e Informatica, Mexico, 2002.

[13] N. Kakwani. On a class of poverty measures. *Econometrica*, 1980.

[14] N. Kakwani. Statistical infrence in the measurement of poverty. *The Review of Economics and Statistics*, pages 632–639, 1993.

[15] B. Kreider and J. Pepper. Inferring disability status from corrupt data. Mimeo, 2004.

[16] Charles Manski. *Partial Identification of Probability Distributions.* Springer-Verlag, first edition, 2003.

[17] Francesca Molinari. Identification of probability distributions with misclassified data. Mimeo, Cornell University, 2003.

[18] M. Ravallion. Poverty rankings using noisy data on living standards. *Economics Letters*, 45:481–485, 1994.

[19] SEDESOL. 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, 2002.

[20] C. Seidl. Poverty measurement: A survey. In D. Boss, M. Rose, and C. Seidl, editors, *Welfare and Efficiency in Public Economics.* Springer-Verlag, 1998.

[21] A. Sen. Poverty: An ordinal approach to measurement. *Econometrica*, 54, 1976.

[22] T. Srinivassan. Data base for development analysis: An overview. *Journal of Development Economics*, 44:3–27, 1994.

[23] S. Stigler. The asymptotic distribution of the trimmed mean. *The Annals of Statistics*, 1:472–477, 1973.

[24] J. Strauss and D. Thomas. Measurement and mismeasurement of social indicators. *American Economic Review*, 86:30–34, 1996.

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