TESTING FOR PRO-POORNESS OF GROWTH, WITH AN APPLICATION TO MEXICO

BY ABDELKRIM ARAAR* AND JEAN-YVES DUCLOS

Département d'économique and CIRPÉE, Université Laval, Canada

MATHIEU AUDET

GRÉDI, Université de Sherbrooke, Canada

PAUL MAKDISSI

Department of Economics, University of Ottawa, Canada

This paper proposes techniques to test for whether growth has been pro-poor. We first review different definitions of pro-poorness and argue for the use of methods that can generate results that are robust over classes of pro-poor measures and ranges of poverty lines. We then provide statistical procedures that rely on the use of sample data to infer whether growth has been pro-poor in a population. We apply these procedures to Mexican household surveys for 1992, 1998, and 2004. We find strong normative and statistical evidence that Mexican growth has been absolutely anti-poor between 1992 and 1998, absolutely pro-poor between 1998 and 2004 and between 1992 and 2004, and relatively pro-poor between 1992 and 2004 and between 1998 and 2004. The relative assessment of the period between 1992 and 1998 is statistically too weak to lead to a robust evaluation of that period.

1. INTRODUCTION

It would seem uncontroversial to define the pro-poorness of growth on the basis of the impact of growth on the poor. Like many other distributive concepts, the precise definition of pro-poorness appears, however, to be essentially a matter of normative judgment.[1]

There are at least three main sources of contention in the literature in trying to make the assessment of pro-poorness operational. The first issue is whether the concept of pro-poorness should be absolute or relative. A second issue is what poverty line should be chosen to separate the poor from the non-poor. A final issue is how to aggregate the impact of growth on the poor. This paper's main contribution is to address all these issues in a normatively "robust" framework, analogous to the stochastic dominance framework used for comparing poverty and

[1]See, among many recent contributions to that debate, Bourguignon (2003), Bruno *et al.* (1998), Dollar and Kraay (2002), Eastwood and Lipton (2001), United Nations (2000), and World Bank (2000).

inequality.[2] An additional contribution of the paper to the literature is to develop procedures to test statistically for the pro-poorness of distributive changes, procedures that help make checks for pro-poor growth "robust" to the presence of sampling variability.

To do this, we first rely on the definitional framework of Duclos (2009). Roughly speaking, and according to that framework, a relative definition of pro-poorness judges a distributive change to be pro-poor if the proportional change in the incomes of the poor is no less than some norm, often set as the growth rate in mean income or in some quantile such as median income.[3] This relative definition is consistent with the approach followed by Kakwani and Pernia (2000) and Son (2004).

For an absolute definition, the incomes of the poor need to grow by an absolute amount that is no less than some norm, often (including in this paper) set to zero or to some function of the absolute change in mean or median incomes. For instance, Ravallion and Chen (2003) suggest that growth should be considered pro-poor if it reduces the Watts index. This implicitly implies an absolute norm equal to 0—growth is pro-poor if it reduces some measure of absolute poverty. These different definitions are analogous to the usual concepts of absolute and relative poverty. With relative poverty, the poverty line is usually defined as a proportion of some central tendency of an income distribution; with absolute poverty, the real level of the poverty line normally remains constant even if the income distribution changes.

Klasen (2008) suggests that an absolute approach may be of interest when a policy analyst is concerned about the pace of poverty reduction, while a relative approach might be more appropriate if the analyst wishes to determine whether the opportunities provided by a growth pattern have helped the poor disproportionately. Essama-Nssah (2005) has also recently proposed a unifying framework that allows for both an absolute and a relative definition through a common measurement system.

The framework proposed in Duclos (2009) differs from the previous literature. He develops a stochastic dominance approach that serves to identify which growth patterns would be deemed pro-poor by a wide spectrum of pro-poor judgments. The framework does this by investigating how pro-poor judgments can be made robust to classes of pro-poor evaluation functions and to ranges of poverty lines. In that sense, the proposed approach enables getting around the difficulty of having to choose (1) a poverty line to separate the poor from the non-poor, and (2) a set of normative weights to differentiate the poor among themselves.[4]

[2]See, for instance, Atkinson (1970), Shorrocks (1983), and Foster and Shorrocks (1988a, 1988b).
[3]Although absolute poverty is usually of greater concern in developing countries, interest in relative poverty has nevertheless gained significant ground in developed economics (see, among many others, Atkinson *et al.*, 2002) and it is also emerging as an important issue in developing countries too. This is either because inequality may be potentially bad for growth (see, for instance, Alesina and Rodrik, 1994; Deininger and Squire, 1998), because it usually makes poverty fall slower for a given level of economic growth, because it breeds relative deprivation, economic isolation and social exclusion, or because it can deemed bad on its own for well-known ethical reasons—such as those developed in Rawls (1971).
[4]Several different approaches have been proposed to separate the poor from the non-poor and to compute indices of pro-poorness. See, for instance, McCulloch and Baulch (1999), Kakwani and Pernia (2000), Ravallion and Datt (2002), Kakwani *et al.* (2003), Ravallion and Chen (2003), Klasen (2004), Son (2004), Essama-Nssah (2005), Essama-Nssah and Lambert (forthcoming), and Kakwani and Son (2008).

In practice, household data surveys are of course needed to check if growth is pro-poor or not. This leads us to consider issues of estimation, sampling variability and statistical inference. We start by proposing sets of null and alternative hypotheses for testing for absolute and relative pro-poorness of growth. We then define various estimators of the statistics of interest and derive their sampling distribution—taking full account of the complexity of the usual forms of sampling design, and of the fact that these statistics of interest typically involve non-linear combinations of estimators from different, though sometimes dependent, samples. This enables us *inter alia* to draw confidence intervals around the differences that must be signed in order to conclude that a change has been robustly pro-poor or not. To our knowledge, this is the first paper to address the issue of pro-poor comparisons within such a general statistical setting.

We then illustrate the usefulness of these techniques by applying them to Mexico's National Income and Expenditure Surveys collected in 1992, 1998, and 2004. Two main conclusions emerge. First, we find strong normative support and statistical evidence that Mexican growth has been absolutely anti-poor between 1992 and 1998, absolutely pro-poor between 1998 and 2004 and between 1992 and 2004, and relatively pro-poor between 1992 and 2004 and between 1998 and 2004. This is useful and robust evidence on Mexico's distributional change during a period of significant economic turbulence and transformations.

Second, we see how and why observing pro-poorness in samples does not imply that we can immediately infer it for populations of interest. For instance, although the sample estimates might naively suggest that the distributive change between 1992 and 1998 is first-order relatively pro-poor, the use of our statistical inference techniques suggests that the evidence is statistically too weak to lead to a robust relative pro-poor evaluation of that period. We also see that, contrary to what measurement theory usually implies for poverty dominance relationships, moving to higher orders does not necessarily lead to more definite comparisons. For instance, statistical uncertainty in comparing 1992 to 1998 is amplified when moving from first-order to second-order relative comparisons, which also means that second-order relative comparisons are statistically less strong than first-order ones.

The rest of the paper is as follows. Section 2 presents the measurement framework. Section 3 deals with issues of estimation and statistical inference. Section 4 applies the measurement and statistical techniques to Mexican data. Section 5 concludes.

2. Measurement Framework

2.1. *The General Setting*

Let $\mathbf{y}_1 = \left(y_1^1, y_2^1, \ldots, y_{n_1}^1 \right) \in \mathfrak{R}_+^{n_1}$ be a vector of non-negative initial incomes[5] (at time 1) of size n_1, and let $\mathbf{y}_2 = \left(y_1^2, y_2^2, \ldots, y_{n_2}^2 \right)$ be an analogous vector of posterior incomes (at time 2) of size n_2.

The following draws extensively from Duclos (2009). To determine whether the movement from $\mathbf{y}_1$ to $\mathbf{y}_2$ is pro-poor, we first need to define a standard with

---

[5]Or consumption, wealth, or any other welfare indicator of interest. See also Klasen (2008) for a discussion of measurement issues in non-income dimensions.

which this assessment can be made. First consider the case of a relative standard, which we will take in this paper as the growth in average incomes, denoted by g.

Intuitively, for growth to be relatively pro-poor, we wish the poor's "representative" income to undergo a proportional change that is no less than g. That growth g can be negative as well as positive. Then denote by $z > 0$ a poverty line, defined in real terms. Let $W(\mathbf{y}_1, \mathbf{y}_2, g, z)$ be the pro-poor evaluation function that we want to use. It is defined as the difference between two evaluation functions $\Pi(\mathbf{y}_1, z)$ and $\Pi^*(\mathbf{y}_2, g, z)$, each for time 1 and time 2, respectively, and which are analogous to poverty indices for each of the two time periods:

(1) $$W(\mathbf{y}_1, \mathbf{y}_2, g, z) \equiv \Pi^*(\mathbf{y}_2, g, z) - \Pi(\mathbf{y}_1, z).$$

The change from $\mathbf{y}_1$ to $\mathbf{y}_2$ will be deemed pro-poor if $W(\mathbf{y}_1, \mathbf{y}_2, g, z) \leq 0$.

Clearly, whether the distributional change will be deemed pro-poor will depend on the way in which z, Π, and Π^* will be chosen. To put some structure on the form of W in which we should be interested, a few axioms can be imposed. The first is a *focus axiom*, through which W is not sensitive to marginal changes in values of $\mathbf{y}_1$ that exceed z. The second is a *population invariance* axiom, which says that adding a replication of a population to that same population has no impact on W. A third axiom is that of *anonymity*: permuting the incomes of any two persons in any given distribution should not affect pro-poor judgments. A fourth *normalization* axiom says that in the absence of distributional change (and also with no change in mean income), then $W = 0$. A fifth *monotonicity* axiom imposes that, for a given g, if anyone's posterior income increases, W should not increase, and may sometimes fall. A sixth *distribution sensitivity* axiom says that the evaluation functions Π should give more weight to the poorer than to the not-so-poor among the poor: shifting incomes from the richer to the poorer is by itself a pro-poor distributional change. This axiom is also known as the Pigou–Dalton principle in the welfare literature.

2.2. *Relative Pro-Poor Judgments*

We can then also axiomatize the *relative* view of pro-poorness. Suppose that $\mathbf{y}/(1 + g) = \mathbf{y}'/(1 + g')$. Formally then, according to the *relative* axiom, $\mathbf{y}$ and $\mathbf{y}'$ should be judged equally pro-poor by W (using standards g and g' respectively) regardless of the initial distribution $\mathbf{y}_1$.

Combined together, the axioms that we have invoked until now define what can be termed to be first-order (if we exclude the *distribution sensitivity* axiom) and second-order (if we include the *distribution sensitivity* axiom) classes of relative pro-poor evaluation functions. Denote those classes as $\Omega^1(g, z^+)$ and $\Omega^2(g, z^+)$, respectively. Most poverty indices belong to both of these classes, including the Watts (1968) index, the second class of indices proposed by Clark *et al.* (1981), and the class of indices proposed by Chakravarty (1983). Important exceptions are the headcount index and the FGT indices (see (2) below) with $0 \leq \alpha < 1$, which belong only to $\Omega^1(g, z^+)$.

Now let $F_j(y)$ be the distribution function of distribution j. Also define as $Q_j(p)$ the quantile function for distribution F_j. This is formally defined as

$Q_j(p) = \inf\{s \geq 0 | F_j(s) \geq p\}$ for $p \in [0, 1]$. With a continuous distribution and a strictly positive income density, $Q(p)$ is simply the usual inverse of the distribution function, and it is the income of that individual who is at rank p in the distribution.

The popular class of FGT indices (Foster, Greer, and Thorbecke, 1984) is then given by:

$$(2) \qquad P_j(z; \alpha) = \int_0^{F_j(z)} \left(1 - Q_j(p)/z\right)^{\alpha} dp.$$

$P_j(z; \alpha = 0)$ is the headcount index (and the distribution function) at z, and $P_j(z; \alpha = 1)$ is the average poverty gap.

Formally, the class $\Omega^1(g, z^+)$ regroups all of the functions W that satisfy the focus, the population invariance, the anonymity, the monotonicity, the normalization and the relative axioms, and for which $z \leq z^+$. Duclos (2009) shows that a movement from $\mathbf{y}_1$ to $\mathbf{y}_2$ will be judged pro-poor by all such functions if and only if

$$(3) \qquad P_2((1+g)z; \alpha = 0) \leq P_1(z; \alpha = 0) \quad \text{for all } z \in \left[0, z^+\right].$$

Verifying (3) simply involves checking whether—over the range of poverty lines $[0, z^+]$—the headcount index in the initial distribution is larger than the headcount index in the posterior distribution when that distribution is normalized by $1 + g$.

An alternative and equivalent way of checking whether a distributional change can be declared first-order relatively pro-poor is to compare the ratio of the quantiles to $(1 + g)$, or, if g is growth in mean income, to compare the growth in quantiles to the growth in the mean. That this, we check whether, for all $p \in [0, F_1(z^+)]$,

$$(4) \qquad \frac{Q_2(p)}{Q_1(p)} \geq 1 + g$$

or whether

$$(5) \qquad \frac{Q_2(p) - Q_1(p)}{Q_1(p)} \geq g.$$

Using (5) is equivalent to Ravallion and Chen's (2003) suggestion to use "growth incidence curves" to check whether growth is pro-poor. These curves show the growth rates of living standards at different ranks in the population.

First-order pro-poor judgments can be demanding in expansion periods. They require *all* quantiles of the poor to undergo a rate of growth at least as large as the rate of growth in mean income. Adding the distribution-sensitivity axiom to the earlier axioms defines a smaller second-order class of relative pro-poor evaluation functions $\Omega^2(g, z^+)$. Formally, $\Omega^2(g, z^+)$ is made of all functions $W(\cdot, \cdot, g, z)$ that satisfy the focus, the population invariance, the anonymity, the monotonicity, the normalization, the distribution sensitivity and the relative axioms, and for which $z \leq z^+$.

It can then be shown that a movement from $\mathbf{y}_1$ to $\mathbf{y}_2$ will be judged pro-poor by all pro-poor evaluation functions $W(\cdot, \cdot, g, z)$ that are members of $\Omega^2(g, z^+)$ if and only if

$$(6) \qquad P_2((1+g)z; \alpha = 1) \leq P_1(z; \alpha = 1) \quad \text{for all } z \in [0, z^+].$$

Verifying (6) simply involves checking whether the average poverty gap in the initial distribution is larger than that in the posterior distribution when that distribution is normalized by $1 + g$ and this, over the range of poverty lines $[0, z^+]$.

Note that although only one value of the FGT index is used ($\alpha = 0$ or $\alpha = 1$) to implement the first-order and second-order tests, if (3) or (6) holds, then the judgment that the change is pro-poor must be upheld by all members of the class of judgments given by $\Omega^1(g, z^+)$ or $\Omega^2(g, z^+)$. These judgments can use any sets of normative weights, so long as these are consistent with the axioms that define the two classes.

As for first-order pro-poor judgments, there are alternative ways of checking condition (6). The cumulative income up to rank p (the Generalized Lorenz curve at p; *see* Shorrocks, 1983) is given by

$$(7) \qquad C_j(p) = \int_0^p Q_j(q)\, dq.$$

The use of the Generalized Lorenz curve provides an intuitive *sufficient* condition for checking second-order relative pro-poorness. A distributional change is indeed second-order relatively pro-poor if for all $p \in [0, F_2((1+g)z^+)]$,

$$(8) \qquad \lambda(p) \equiv \frac{C_2(p)}{C_1(p)} \geq 1 + g.$$

Expression (8) involves computing the growth rates in the cumulative incomes of proportions p of the poorest, and to compare[6] those growth rates to g. For $1 + g$ equal to the ratio of mean income, condition (8) is equivalent to checking whether the Lorenz curve for $\mathbf{y}_2$ is above that of $\mathbf{y}_1$ for the range of $p \in [0, F_2((1+g)z^+)]$.

2.3. *Absolute Pro-Poor Judgments*

Absolute pro-poor judgments are made by comparing the absolute change in the poor's incomes to some absolute pro-poor standard. Denote that standard as *a*. The axiom of *absolute* pro-poorness says essentially that Π^* should be "translation invariant" in $\mathbf{y}$ and *a*, or that the pro-poor judgment should be neutral whenever the poor gain in absolute terms the same as the standard *a*. This reference point is consistent with the view that a change is good for the poor if it increases the poor's absolute living standards (e.g., Ravallion and Chen, 2003). Hence, the absolute axiom demands that if $\mathbf{y} + a = \mathbf{y}'$, then $W(\mathbf{y}, \mathbf{y}', a, z) = 0$.

---

[6]This is similar to a condition provided by Son (2004), with the difference that (8) is there checked over all $p \in [0, 1]$.

This allows us to define formally the class of first-order absolute pro-poor evaluation functions $\tilde{\Omega}^1(a, z^+)$ as made of all those functions $W(\cdot, \cdot, a, z)$ that satisfy the focus, the population, the anonymity, the monotonicity, the normalization and the absolute axioms, and for which $z \leq z^+$. We will later set a to zero for the empirical illustration of this paper. This corresponds to what Klasen (2008) terms "weak absolute pro-poor growth."[7]

It can then be shown that a movement from $\mathbf{y}_1$ to $\mathbf{y}_2$ will be judged first-order absolutely pro-poor (that is, pro-poor by all evaluation functions $W(\cdot, \cdot, a, z)$ that are members of $\tilde{\Omega}^1(a, z^+)$ if and only if

$$(9) \qquad P_2(z + a; \alpha = 0) \leq P_1(z; \alpha = 0) \text{ for all } z \in [0, z^+].$$

An equivalent way of checking whether a distributional change can be declared first-order absolutely pro-poor is to compare the absolute change in the values of the quantiles for all $p \in [0, F_1(z^+)]$:

$$(10) \qquad Q_2(p) - Q_1(p) \geq a.$$

An analogous result holds for absolute second-order pro-poor judgments. These judgments are made on the basis of the class of indices $\tilde{\Omega}^2(a, z^+)$, which is defined as for $\tilde{\Omega}^1(a, z^+)$ but with the additional requirement of distribution sensitivity. A movement from $\mathbf{y}_1$ to $\mathbf{y}_2$ is then judged to be second-order absolutely pro-poor if and only if

$$(11) \qquad (z + a)P_2((z + a; \alpha = 1) \leq zP_1(z; \alpha = 1) \text{ for all } z \in [0, z^+].$$

A sufficient condition for condition (11) is then to verify whether, for all $p \in [0, F_2(z^+ + a)]$, the change in the average income of the bottom p proportion of the population is larger than a:

$$(12) \qquad \frac{C_2(p) - C_1(p)}{p} \geq a.$$

3. Estimation and Statistical Inference

In practice, household data surveys are needed to check if growth is pro-poor or not. This forces us to deal with issues of estimation, sampling variability and statistical inference. Indeed, a difference in some statistics observed across samples may not be empirically strong enough to be significant from a statistical point of view. Although this is a point whose importance is increasingly well understood in the applied poverty measurement literature, it raises methodological issues that are not straightforward in the context of this paper. These issues deal with the fact that we are interested here in (dominance) tests over ranges of parameter values, and in

[7]This is also consistent with Ravallion and Chen (2003). Testing for what Klasen (2008) defines "strong pro-poor growth" (equating a to the "mean absolute improvement of the non-poor") would certainly have led to different results—see also Duclos et al. (2009).

statistics that are non-linear functions of estimators estimated across different (dependent or independent) samples. We consider these issues in turn.

3.1. *Null and Alternative Hypotheses for Testing Pro-Poorness of Growth*

Each of the conditions noted in Sections 2.2 and 2.3 takes the form of testing whether $\Delta^s(z) \leq 0$ or $\Delta^s(p) \geq 0$ over some *range* of z or p. This therefore involves testing *jointly* over a *set* of null hypotheses. For primal tests of pro-poorness (those using estimators based on monetary thresholds z), our formulation of our null hypothesis is thus that of a union of null hypotheses,

$$(13) \qquad H_0 : \Delta^s(z) > 0 \text{ for some } z \in [0, z^+],$$

to be tested against an alternative hypothesis that is an intersection of alternative hypotheses,

$$(14) \qquad H_1 : \Delta^s(z) \leq 0 \text{ for all } z \in [0, z^+].$$

For dual tests (those that use estimators based on percentiles p), we formulate a union of null hypotheses,

$$(15) \qquad H_0 : \Delta^s(p) < 0 \text{ for some } p \in [0, \mathbb{F}(z^+)],$$

to be tested against an intersection of alternative hypotheses,

$$(16) \qquad H_1 : \Delta^s(p) \geq 0 \text{ for all } p \in [0, \mathbb{F}(z^+)],$$

where $\mathbb{F}(z^+)$ is some function of z^+. Our decision rule is to reject the union set of null hypotheses in favor of the intersection set of alternative hypotheses only if we can reject each of the individual hypotheses in the null set at a $100 \cdot \theta\%$ significance level. This can be conveniently done by drawing $100 \cdot (1 - \theta)\%$ one-sided confidence intervals, a devise we use repeatedly in the empirical application below.

To see this more precisely, denote by $\hat{\Delta}^s(z)$ the sample estimator of $\Delta^s(z)$, by $\Delta_0^s(z)$ its sample value, and by $\sigma_{\hat{\Delta}^s(z)}^2$ the sampling variance of $\hat{\Delta}^s(z)$. Let $\zeta(\theta)$ be the $(1 - \theta)$-quantile of the normal distribution. Given that by the law of large numbers and the central limit theorem, all of the estimators used in this paper can be shown to be consistent and asymptotically normally distributed, we can use $\Delta_0^s(z) + \sigma_{\hat{\Delta}^s(z)}\zeta(\theta)$ as the upper bound for a one-sided confidence interval for $\Delta^s(z)$. A confidence interval that is upper bounded by $\Delta_0^s(z) + \sigma_{\hat{\Delta}^s(z)}\zeta(\theta)$ shows all of the values of η for which we could not reject a null hypothesis $H_0 : \Delta^s(z) > \eta$ in favor of $H_1 : \Delta^s(z) \leq \eta$. Setting $\eta = 0$ (see (13) and (14)), our decision rule is then to reject the set of null hypotheses (13) in favor of (14) if:

$$(17) \qquad \Delta_0^s(z) + \sigma_{\hat{\Delta}^s(z)}\zeta(\theta) < 0 \,\forall z \in [0, z^+].$$

For dual tests, we proceed similarly to (17), noting that the signs in (15) and (16) are inverted. We thus build a confidence interval $\Delta_0^s(p) - \sigma_{\hat{\Delta}^s(p)}\zeta(\theta)$ and reject (15) in favor of (16) if

$$(18) \qquad \Delta_0^s(p) - \sigma_{\hat{\Delta}^s(p)}\zeta(\theta) > 0 \, \forall p \in \left[0, \mathbb{F}(z^+)\right]$$

for some distribution function $\mathbb{F}$.

3.2. *Estimation and Sampling Variability*

We now need to define and assess $\hat{\Delta}^s(z)$, $\hat{\Delta}^s(p)$, $\sigma_{\hat{\Delta}^s(z)}$ and $\sigma_{\hat{\Delta}^s(p)}$. As we will see, these statistics are non-linear functions of estimators estimated across different (dependent or independent) samples, and for that reason estimating their sampling variability raises difficulties that do not arise for simpler statistics (such as for FGT indices with deterministic poverty lines).

Let N_j be a number of independently and identically distributed sample observations of incomes drawn from a distribution $j, y_j^1, \ldots, y_j^N$. Denoting $f_+ = \max(f, 0)$, a natural estimator of the FGT index $P_j((1 + g)z; \alpha)$ is given by

$$(19) \qquad \hat{P}_j((1+\hat{g})z; \alpha) = N_j^{-1} \sum_{h=1}^{N_j} \left(\frac{(1+\hat{g})z - y_j^h}{(1+\hat{g})z} \right)_+^\alpha$$

where

$$(20) \qquad 1 + \hat{g} = \hat{\mu}_2 / \hat{\mu}_1$$

and

$$(21) \qquad \hat{\mu}_j = N_j^{-1} \sum_{h=1}^{N_j} y_j^h.$$

The p-quantile $Q_j(p)$ is estimated as

$$(22) \qquad \hat{Q}_j(p) = \min\left(y \,\middle|\, \hat{F}_j(y) \ge p\right),$$

and the empirical distribution function is given by

$$(23) \qquad \hat{F}_j(y) = \hat{P}_j(z; 0).$$

Note that there are two sources of sampling variability in (19): the first comes from summing over random sample observations (assuming g to be known),

$$(24) \qquad \hat{P}_j((1+g)z; \alpha) = N_j^{-1} \sum_{h=1}^{N_j} \left(\frac{(1+g)z - y_j^h}{(1+g)z} \right)_+^\alpha,$$

and the second comes from sampling variability in the estimator $\hat{g}$. Using a first-order approximation to (19), we find

$$
\begin{aligned}
\hat{P}_2((1+\hat{g})z; \alpha) - P_2((1+g)z; \alpha) &= \hat{P}_2((1+g)z; \alpha) - P_2((1+g)z; \alpha) \\
&\quad + ((1+g)z)P_2^{gz}((1+g)z; \alpha)\left(\frac{\hat{\mu}_2 - \mu_2}{\mu_2} - \frac{\hat{\mu}_1 - \mu_1}{\mu_1}\right) \\
&\quad + o\left(N_2^{-1/2}\right)
\end{aligned}
$$

(25)

where

(26) $\qquad P_2^{gz}((1+g)z; \alpha) = \alpha((1+g)z)^{-1}(P_2((1+g)z; \alpha-1) - P_2((1+g)z; \alpha))$

for $\alpha > 0$ and

(27) $\qquad\qquad\qquad P_2^{gz}((1+g)z; 0) = f_2((1+g)z) > 0$

(the density at $(1+g)z$) for $\alpha = 0$. $P_2^{gz}((1+g)z; \alpha)$ can be estimated from the samples and is proportional to the effect of $\hat{g}$'s sampling variability on the sampling variability of $\hat{P}_2((1+\hat{g})z; \alpha)$. The small order term $o\left(N_2^{-1/2}\right)$ decreases at a rate faster than $N_2^{1/2}$.

Therefore, we can express $\hat{\Delta}^s(z) - \Delta^s(z)$ as

(28) $\qquad\qquad\qquad \hat{\Delta}^s(z) - \Delta^s(z) = \hat{A} - \hat{B} + o\left(N^{-1/2}\right)$

where

(29) $\quad \hat{A} = \left(\hat{P}_2((1+g)z; \alpha) - P_2((1+g)z; \alpha)\right) + ((1+g)z)P_2^{gz}((1+g)z; \alpha)\left(\frac{\hat{\mu}_2 - \mu_2}{\mu_2}\right)$

and

(30) $\qquad \hat{B} = \left(\hat{P}_1(z; \alpha) - P_1(z; \alpha)\right) - ((1+g)z)P_2^{gz}((1+g)z; \alpha)\left(\frac{\hat{\mu}_1 - \mu_1}{\mu_1}\right).$

The small order term $o\left(N_j^{-1/2}\right)$ in (28) decreases at a rate that is faster than the rates, $N_1^{1/2}$ and $N_2^{1/2}$, at which $\hat{A}$ and $\hat{B}$ decrease as N_1 and N_2 increase towards infinity, so we can ignore it asymptotically. By (20) and (24), each of $\hat{A}$ and $\hat{B}$ in (29) and (30) is a sum of independently and identically distributed (iid) sample observations.

Suppose that the two empirical distributions also come from independent samples, namely, the selection of the sampling units was made independently in each sample. With N_1 and N_2 tending to infinity, we then have (recall (17)):

(31) $\qquad\qquad\qquad \sigma_{\hat{\Delta}^s(z)}^2 = \text{var}\left(\hat{\Delta}^s(z) - \Delta^s(z)\right) \cong \text{var}\left(\hat{A}\right) + \text{var}\left(\hat{B}\right).$

If, however, the two samples are dependent because they come, for instance, from the same panel data, then the asymptotic variance must be estimated jointly over the two samples, and we then have

$$(32) \qquad \sigma^2_{\hat{\Delta}^s(z)} = \text{var}\left(\hat{\Delta}^s(z) - \Delta^s(z)\right) \cong \text{var}\left(\hat{A}\right) + \text{var}\left(\hat{B}\right) - 2\text{cov}\left(\hat{A}, \hat{B}\right).$$

The Appendix provides details on how to compute the terms in (31) and (32).

For the dual (or percentile) approach, we can express

$$(33) \qquad \hat{\Delta}^1(p) - \Delta^1(p) = \left(\frac{\hat{Q}_2(p)}{\hat{Q}_1(p)} - \frac{\hat{\mu}_2}{\hat{\mu}_1}\right) - \left(\frac{Q_2(p)}{Q_1(p)} - \frac{\mu_2}{\mu_1}\right)$$

for first-order relative pro-poorness and

$$(34) \qquad \hat{\Delta}^2(p) - \Delta^2(p) = \left(\frac{\hat{C}_2(p)}{\hat{C}_1(p)} - \frac{\hat{\mu}_2}{\hat{\mu}_1}\right) - \left(\frac{C_2(p)}{C_1(p)} - \frac{\mu_2}{\mu_1}\right)$$

for second-order relative pro-poorness. These expressions can be linearized around the true values $Q_j(p)$, $C_j(p)$ and μ_j using the method of Rao (1973)—see also Duclos and Araar (2006, pp. 290–1). $\hat{\Delta}^1(p) - \Delta^1(p)$ and $\hat{\Delta}^2(p) - \Delta^2(p)$ can then be expressed asymptotically as a weighted sum of the sampling error on the quantile estimators, $\hat{Q}_j(p) - Q_j(p)$, of the sampling error on the estimators of their cumulative function until percentile p, $\hat{C}_j(p) - C_j(p)$, and of the sampling errors on the estimators of the means, $\hat{\mu}_j - \mu_j$. Omitting the subscript j, it can also be shown that

$$(35) \qquad \hat{Q}(p) - Q(p) = N^{-1}\frac{\sum\left(I[y^h < Q(p)] - p\right)}{f(Q(p))} + o\left(N^{-1/2}\right)$$

and

(36)
$$\hat{C}(p) - C(p) = N^{-1}\sum\left\{\left(I[y^h < Q(p)] - p\right)Q(p) + y^h I[y^h < Q(p)] - C(p)\right\} + o\left(N^{-1/2}\right),$$

where $I[y^h < Q(p)]$ is an indicator function taking the value of 1 if its argument is true and 0 otherwise—see Bahadur (1966) and Davidson and Duclos (1997). Again, we can asymptotically ignore the $o(N^{-1/2})$ terms. Noting that $\hat{\mu} - \mu = N^{-1}\Sigma_{h=1}^N y^h - \mu$, all of the expressions $\hat{Q}(p)$, $\hat{C}(p)$ and $\hat{\mu}$ can therefore be expressed as sums of iid variables whose asymptotic sampling distribution can be estimated in a manner similar to the one described for $\hat{A}$ and $\hat{B}$ in the Appendix.

The Appendix shows how sampling design can also be taken into account in making these computations.[8]

---

[8]Demombynes and Hoogeveen (2007) provide confidence intervals around the estimates of $Q_2(p)/Q_1(p) - 1$ (generated by splitting distributions into 15 bins), but it is not entirely clear from the text how these intervals were computed.

4. Has the Mexican Economy Been Pro-Poor in the Last Two Decades?

We apply the above normative and statistical methodology using Mexican data spanning the last decade and a half. Mexico is a particularly interesting economy over which to test the pro-poorness of growth. Mexico has indeed undergone very significant economic changes since 1990 and it also has seen substantial shifts in inequality and poverty during that period.[9] Szekely (2005b) reports, for instance, a sharp increase in poverty around 1996, probably caused by the 1994–1995 economic crisis. After the crisis, which culminated in an important devaluation of Mexico's currency and was probably the most severe in the country's economic history, rapid growth in exports[10] as well as macroeconomic and public sector restructuring led, however, to strong overall economic growth. Recent institutional changes have also encouraged competition and growth in transportation, telecommunications, and power generation and distribution. Clearly, an important issue is whether this period of economic turbulence has been pro-poor, and whether the relatively recent tidal growth has "lifted all boats."

The data used for our application come from the National Income and Expenditure (ENIGH) Surveys collected in 1992, 1998, and 2004. These sample data are representative at the national level. The objective of the ENIGH surveys has been to collect information on incomes and expenditures, goods and services used for self-consumption, and socioeconomic characteristics and labor market activities of all household members. The sampling process was stratified and multi-staged, with the final sampling units being households and all their members.

As is common in Mexico, we use total household income *per capita* as the measure of living standards for all members of a household. Income includes incomes from all sources plus the value of own-produced goods, state benefits, imputed income, gifts, and remittances. To adjust for temporal variation in prices, we express incomes in reference to the 2004 consumer price index. To correct for spatial variation in prices, we assess all incomes in reference to rural prices. This is done by multiplying urban household incomes by the ratio of the rural to the urban poverty line. The rural poverty line in 2004 is often estimated to be around 550 pesos per month *per capita*. We use the product of household size and household sampling weight as an expansion factor to ensure that our samples are representative of the national distribution of the living standards of individuals. For the estimation of standard errors and thus for statistical inference, we take into account the stratification and multi-stage structure of the survey design as explained in Section 3.2.

We begin our investigation by considering the evolution of the density of *per capita* incomes in Figure 1. The distribution of per *capita income* has worsened between 1992 and 1998 since the density curves have shifted to the left. It has

[9]Tétreault (2006) argues that poverty has slightly increased in Mexico between 1982 and 2002. Although Szekely (2005a) reports that there seems to have been no significant trend in poverty and inequality across that period, he finds that there did occur significant cycles of rises and falls in poverty.
[10]This export growth was facilitated by the North American Free Trade and other trade agreements—see, for instance, Nicita (2004).
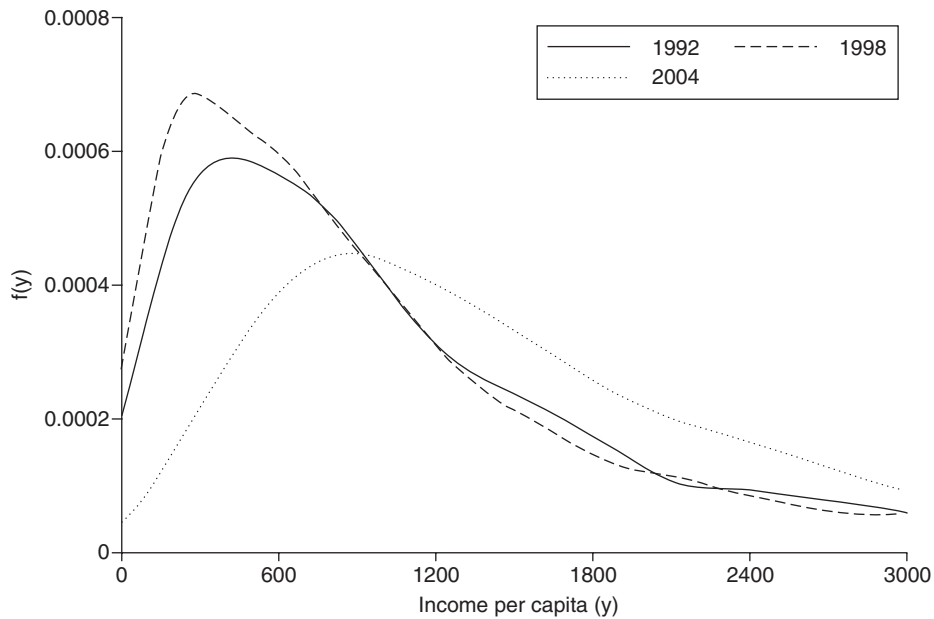
Figure 1. Density Functions

however exhibited a strong and quick recovery between 1998 and 2004, as shown by the shift of the density curve to the right. The estimates of the Lorenz curves and Gini indices presented in Figure 2 and Table 1, respectively, suggest that inequality has decreased between 1992 and 2004. Figures 3 and 4 and the results of Table 1 suggest that absolute poverty, as measured by the headcount and poverty gap indices, has increased between 1992 and 1998 and decreased between 1998 and 2004.

Table 1 also shows the sampling design effect for the estimation of average income *per capita*. The design effect is the ratio of the sampling-design-based estimate of the sampling variance (recall equation (47)) over the estimate of the sampling variance under the assumption of simple random sampling. For average income *per capita*, this stands to between 6.5 and 10.9. The ratio of standard errors is therefore between around 2.5 and 3.3, which also suggests that it is quite important in our case to take into account sampling design in assessing sampling variability—otherwise, we would risk being imprudent in making inferences.

Formal statistical testing for first-order absolute pro-poorness of Mexican growth can be done using the information presented in Figures 5–10. The top line of Figure 5 shows the sample estimates of

$$(37) \qquad \Delta^1(z) = P_{1998}(z; \alpha = 0) - P_{1992}(z; \alpha = 0)$$

for the difference between 1998 and 1992, whereas the dotted bottom curve is the lower bound of the one-sided confidence interval,
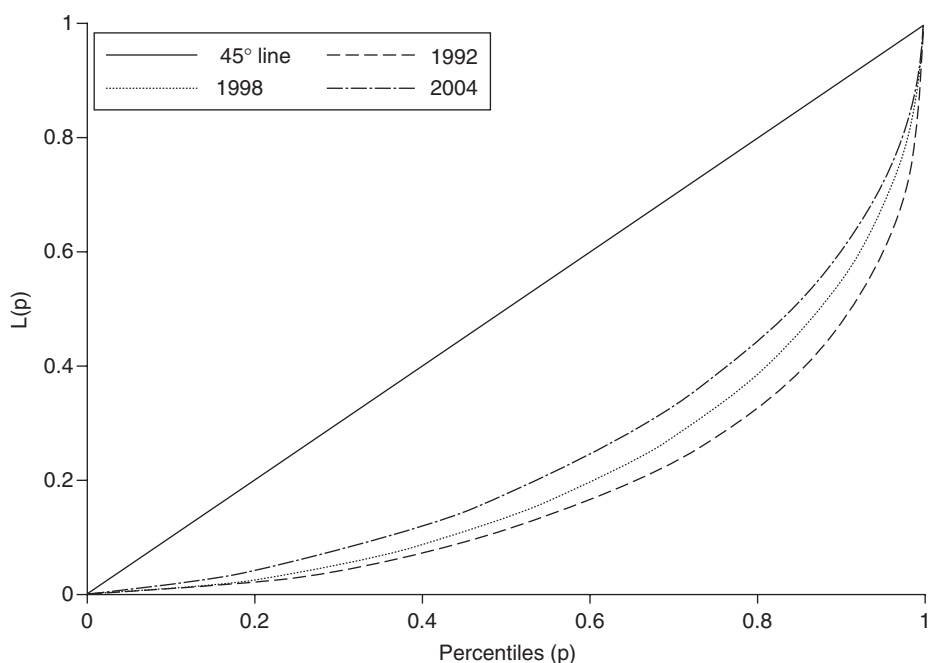
Figure 2. Lorenz Curves

TABLE 1
DESCRIPTIVE STATISTICS FOR ENIGH MEXICAN SAMPLES

| | Year | | |
Statistics	1992	1998	2004
Sample size (households)	10,530	10,952	20,595
Gini index	0.622 (0.017)	0.577 (0.011)	0.524 (0.010)
Average monthly income (μ) per capita	2,115 (142) [6.5]	1,492 (69) [6.8]	2,430 (87) [10.9]
μ_{year}/μ_{1992}	1.000 (0.000)	0.705 (0.057)	1.149 (0.087)
Headcount index ($z = 550$)	0.291 (0.014)	0.354 (0.022)	0.106 (0.009)
Average poverty gap ($z = 550$)	0.129 (0.009)	0.169 (0.015)	0.036 (0.004)

Notes: Standard errors in parentheses.
Sampling design effect in square brackets.

$$(38) \qquad \Delta_0^1(z) - \sigma_{\hat{\Delta}^s_{(z)}} \zeta(\theta).$$

Since $\Delta_0^1(z) - \sigma_{\hat{\Delta}^s_{(z)}} \zeta(\theta) > 0$ is verified in Figure 5 for all reasonable poverty lines, we can infer from our data that growth was absolutely anti-poor during the period 1992 and 1998. The same result obtains from Figure 6 using differences in quantiles. The sample estimates of the difference in quantiles between 1992 and 1998 is shown by the dashed curve, and the upper bound of a one-sided confidence
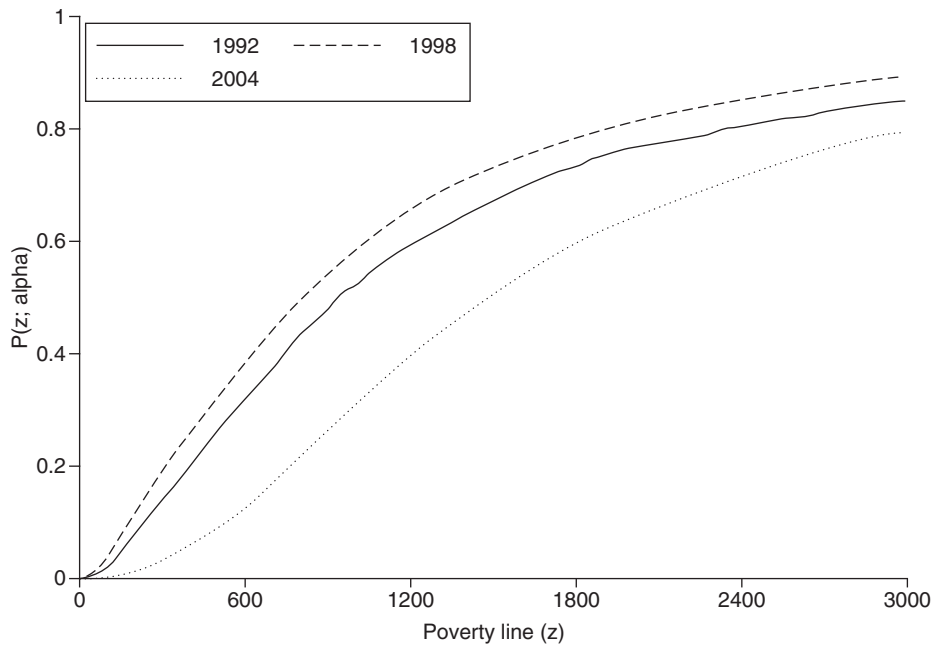
Figure 3. Poverty Headcount Curves: $P(z; \alpha = 0)$ for a range of z
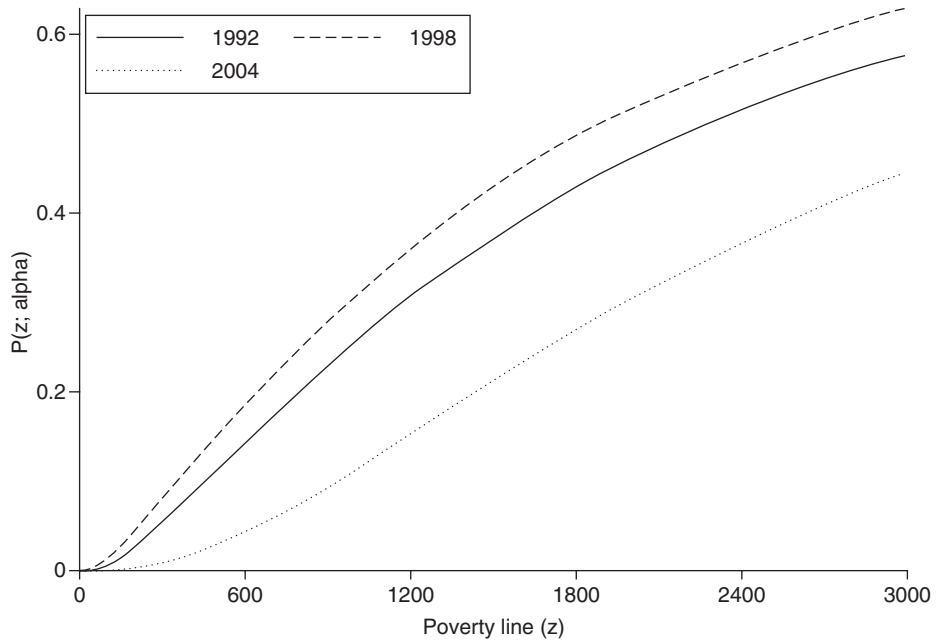
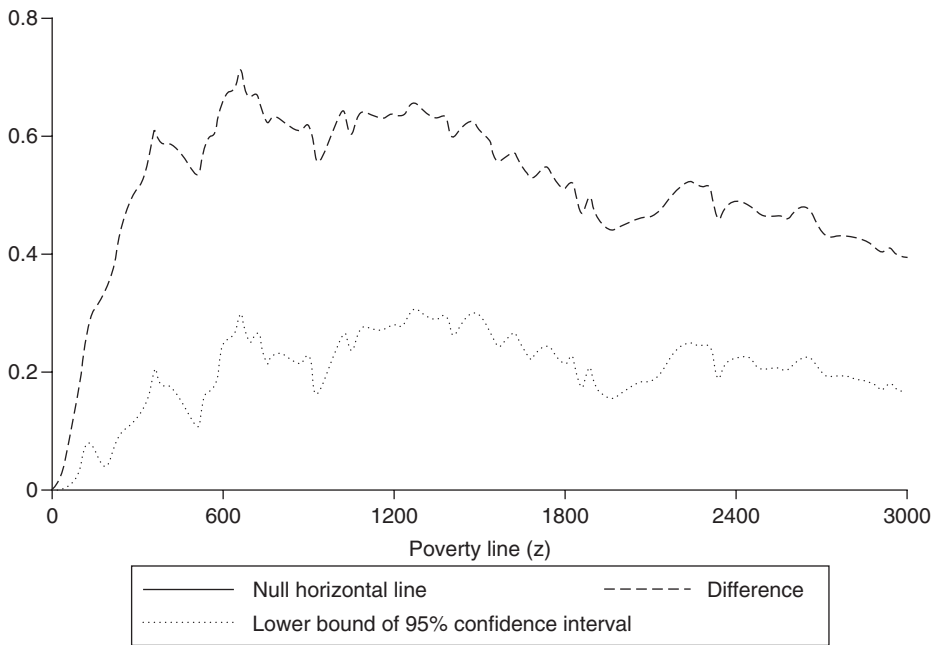Figure 4. Average Poverty Gap Curves: $P(z; \alpha = 1)$ for a range of z

Figure 5. 1992 to 1998 is First-Order Absolutely Anti-Poor. Difference $P_{1998}(z; \alpha = 0) - P_{1992}(z; \alpha = 0)$ is shown on the vertical axis
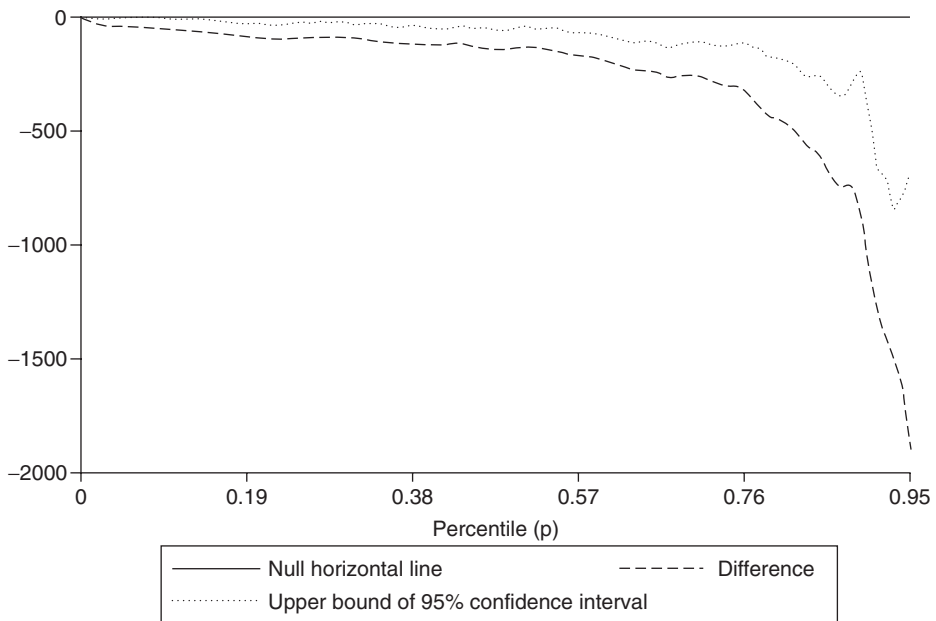

Figure 6. 1992 to 1998 is First-Order Absolutely Anti-Poor. Difference $Q_{1998}(p) - Q_{1992}(p)$ is shown on the vertical axis
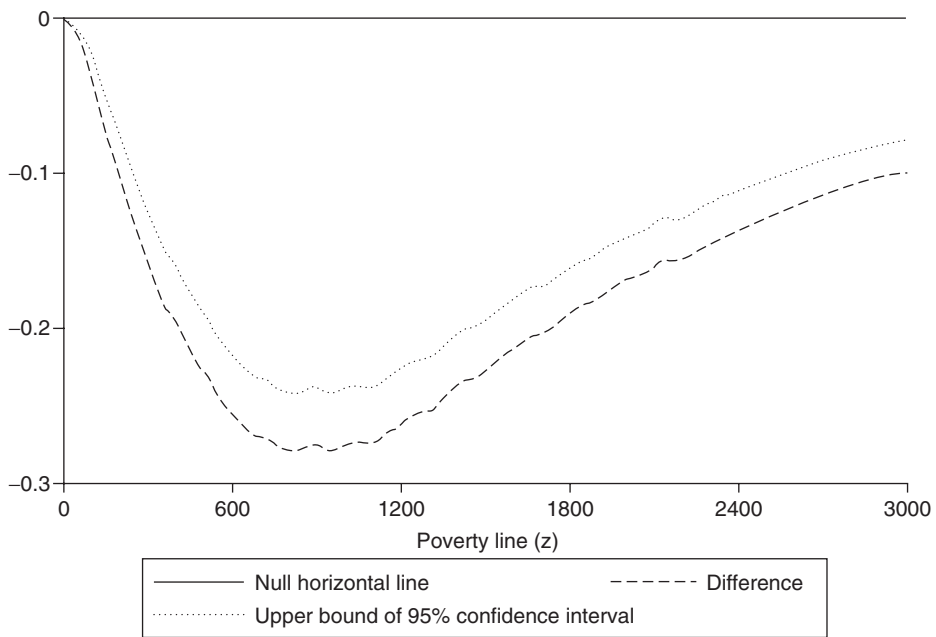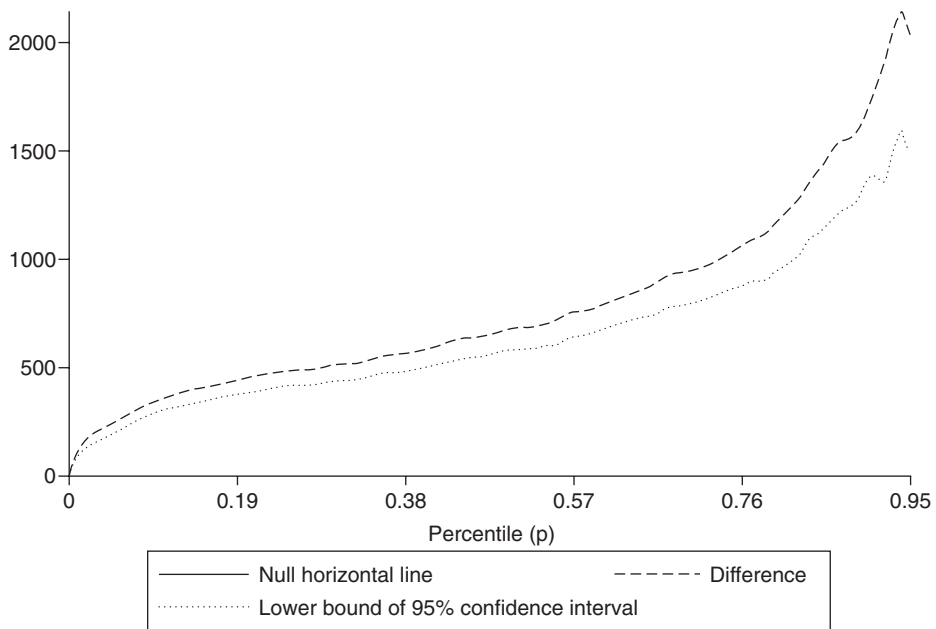
Figure 7. 1998 to 2004 is First-Order Absolutely Pro-Poor. Difference $P_{2004}(z; \alpha = 0) - P_{1998}(z; \alpha = 0)$ is shown on the vertical axis

Figure 8. 1998 to 2004 is First-Order Absolutely Pro-Poor. Difference $Q_{2004}(p) - Q_{1998}(p)$ is shown on the vertical axis
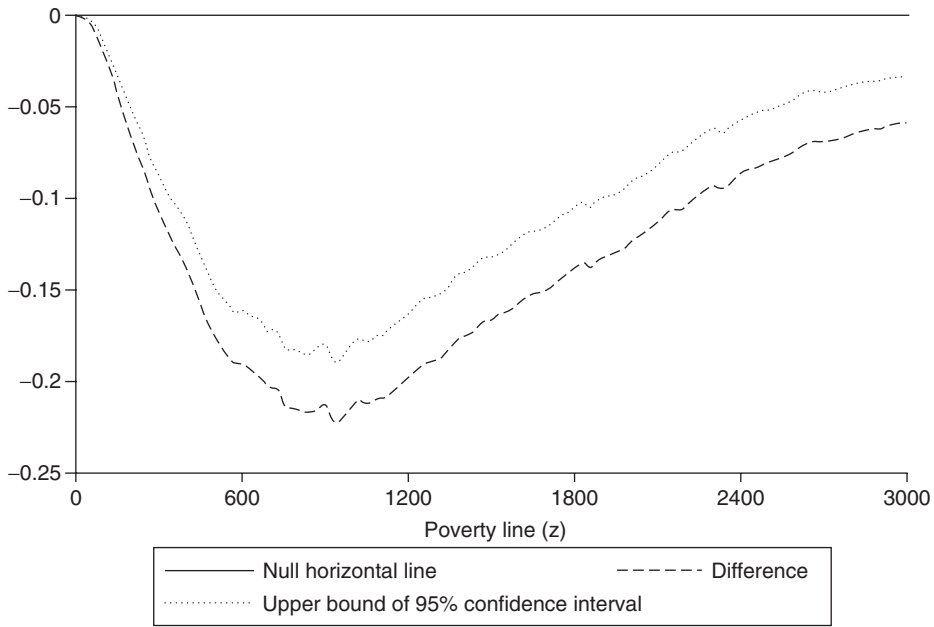
Figure 9. 1992 to 2004 is First-Order Absolutely Pro-Poor. Difference $P_{2004}(z; \alpha = 0) - P_{1992}(z; \alpha = 0)$ is shown on the vertical axis
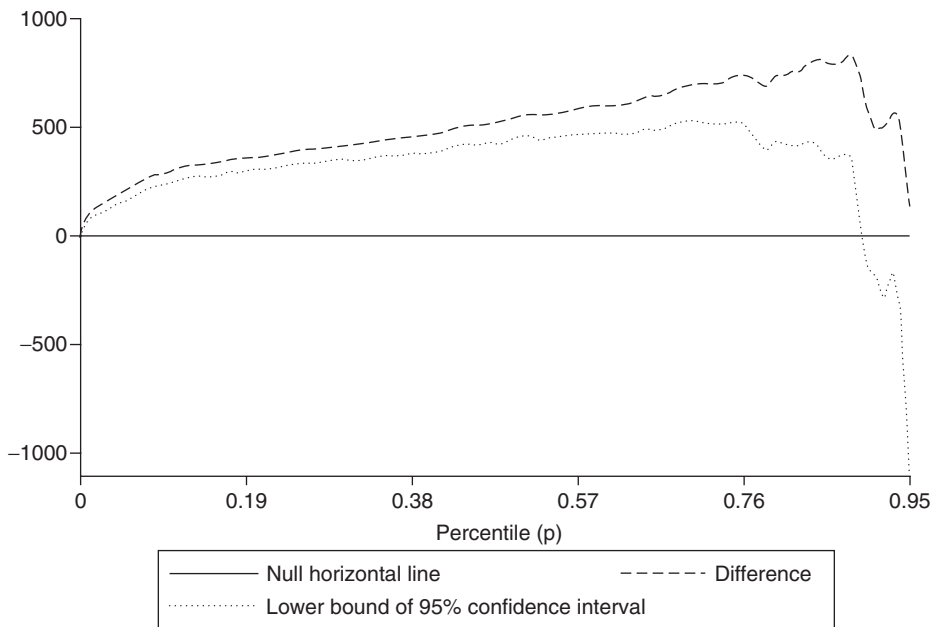

Figure 10. 1992 to 2004 is First-Order Absolutely Pro-Poor. Difference $Q_{2004}(p) - Q_{1992}(p)$ is shown on the vertical axis

interval is shown by the dotted curve. Since we can see in Figure 6 that $\Delta_0^1(p) + \sigma_{\hat{\Delta}^s(p)} \zeta(\theta) < 0$ for all percentiles p between 0 and 0.95, we can conclude from our data that growth was absolutely anti-poor during the period 1992 and 1998.

Opposite results are obtained when comparing 1998 to 2004. Judging from Figures 7 and 8, the change in distribution was first-order absolutely pro-poor. The upper bound of the confidence interval for $\Delta^1(z)$ is everywhere negative, whatever reasonable poverty line is selected (remember that the official rural poverty line is 550 pesos per month *per capita*), and the lower bound of the confidence interval for $\Delta^1(p)$ is everywhere positive, whatever reasonable percentile is selected.

Given the conflicting results reported above, it would seem useful to check for pro-poorness over the entire period 1992 to 2004. This can be done using Figures 9 and 10. The distributive change was almost certainly first-order absolutely pro-poor. The lower bound of the confidence interval for $P_{2004}(z; \alpha = 0) - P_{1992}(z; \alpha = 0)$ is everywhere negative, again whatever reasonable poverty line is selected, and the lower bound of the confidence interval for the difference in quantiles, $Q_{2004}(p) - Q_{1992}(p)$, is everywhere positive, until at least the 0.8 percentile. Thus, the anti-poor movement of 1992 to 1998 was outdone by the pro-poor movement of 1998 to 2004 so that the entire period of 1992 to 2004 can be inferred to be overall first-order absolutely pro-poor.

Given the robust results obtained for first-order pro-poorness, it is not useful to test for second-order pro-poorness since first-order pro-poorness implies second-order pro-poorness. This can be seen by noting that

$$(39) \qquad P_j(z; \alpha = 1) = \int_0^z P_j(y; \alpha = 0)\, dy.$$

If first-order pro-poorness obtains at order 1, then by (39) second-order pro-poorness also obtains. The same relation is obtained by noting from equations (7), (10), and (12) that the Generalized Lorenz curve condition is implied by the quantile condition.

Testing for first-order relative pro-poorness can be done using Figures 11–20. Figure 11 shows why observing pro-poorness in samples does not mean that we can infer it in populations; to go from sample pro-poorness to population pro-poorness, we need to apply statistical inference methods. To see this, note that although average income fell by about 30 percent between 1992 and 1998, the sample estimates of $P_{1998}((1 + g)z;\ \alpha = 0) - P_{1992}(z;\ \alpha = 0)$ suggest that the distributive movement during that period is first-order relatively pro-poor since that difference is always negative in the samples observed. But drawing a confidence interval around the sample estimates make it clear in Figure 11 that the observed differences $P_{1998}((1 + g)z;\ \alpha = 0) - P_{1992}(z;\ \alpha = 0)$ are not statistically significant over a wide range of bottom poverty lines—the upper bounds of the one-sided confidence intervals extend above the zero line for z up to around 600 pesos and p up to around 0.28. Hence, with a conventional level 95 percent of statistical confidence, the first-order relative pro-poor condition is not satisfied.
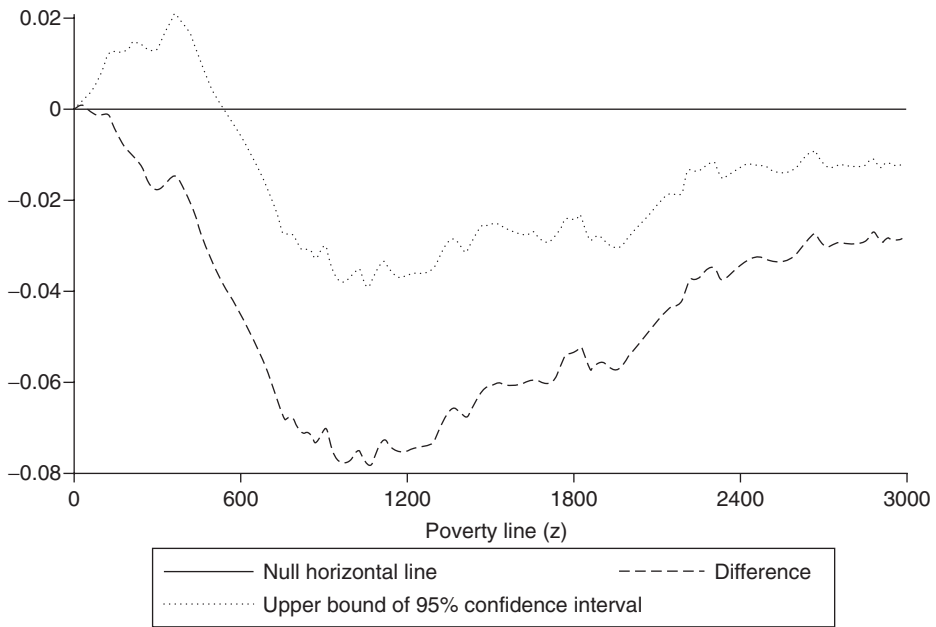
Figure 11. 1992 to 1998 is Not Statistically First-Order Relatively Pro-Poor. Difference $P_{1998}((1+g)z; \alpha = 0) - P_{1992}(z; \alpha = 0)$ is shown on the vertical axis
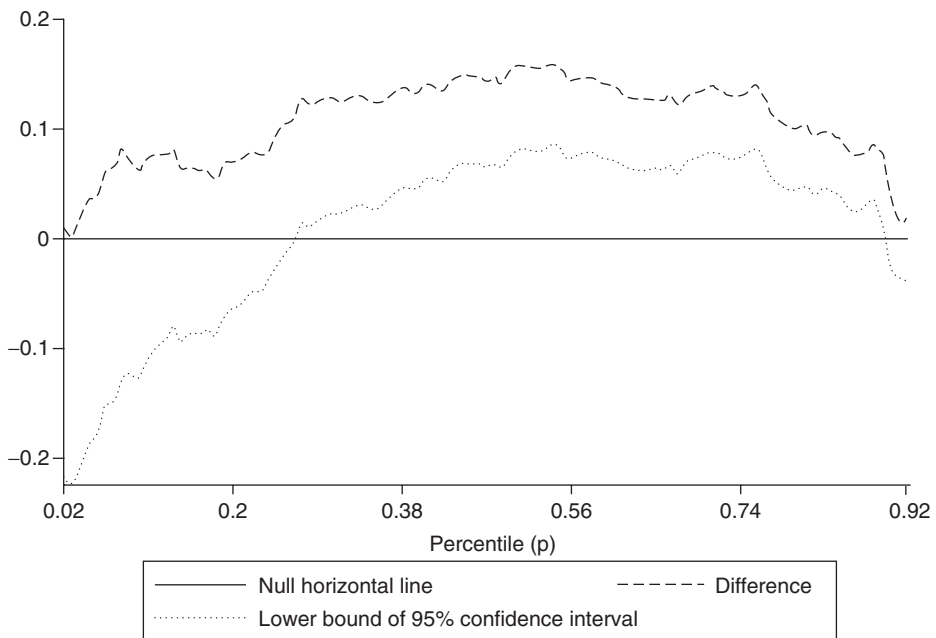

Figure 12. 1992 to 1998 is Not Statistically First-Order Relatively Pro-Poor. Difference $Q_{1998}(p)/Q_{1992}(p) - \mu_{1998}/\mu_{1992}$ is shown on the vertical axis
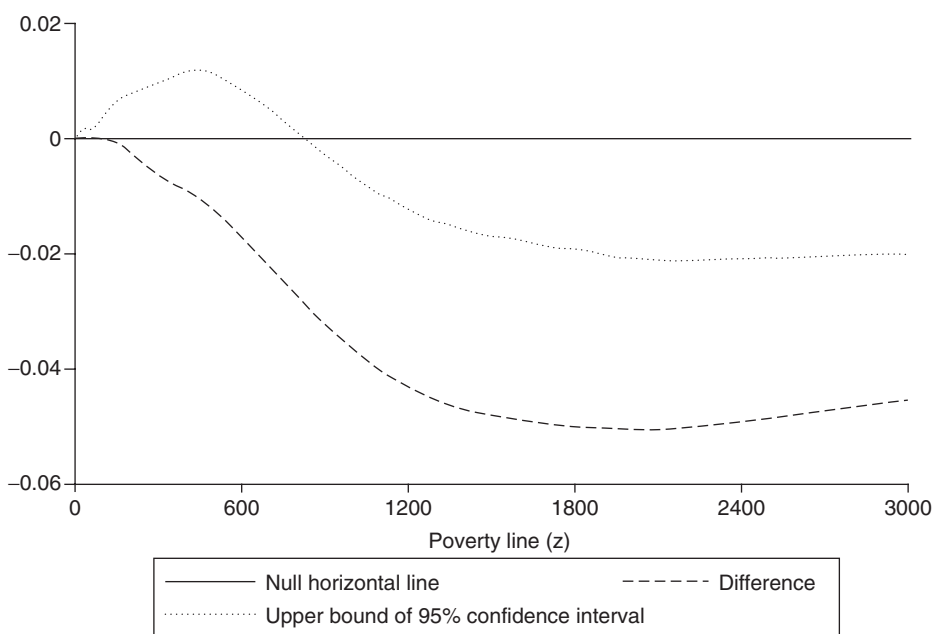
Figure 13. 1992 to 1998 is Not Statistically Second-Order Relatively Pro-Poor. Difference $P_{1998}((1 + g)z; \alpha = 1) - P_{1992}(z; \alpha = 1)$ is shown on the vertical axis
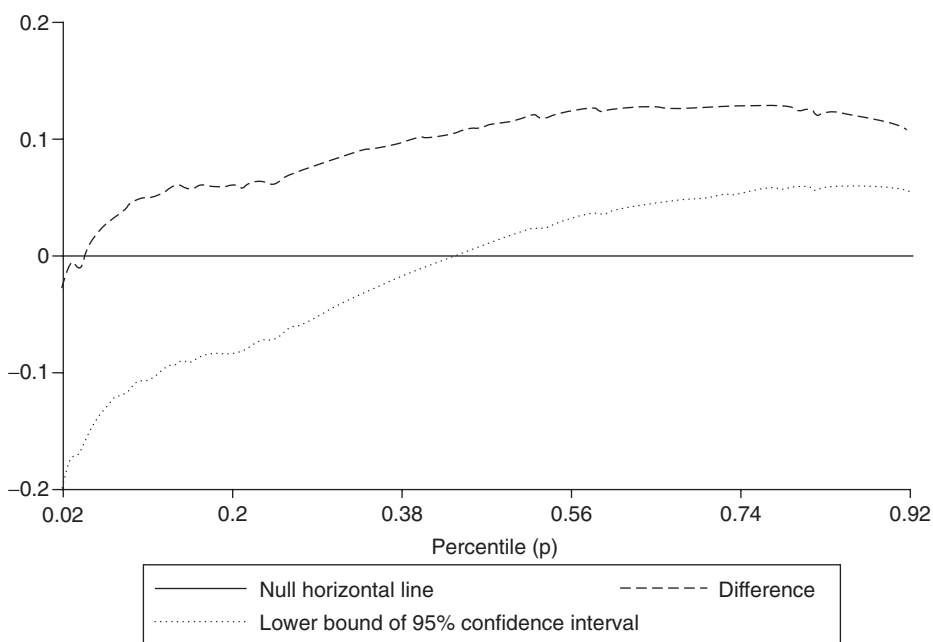

Figure 14. 1992 to 1998 is Not Statistically Second-Order Relatively Pro-Poor. Difference $C_{1998}(p)/C_{1992}(p) - \mu_{1998}/\mu_{1992}$ is shown on the vertical axis
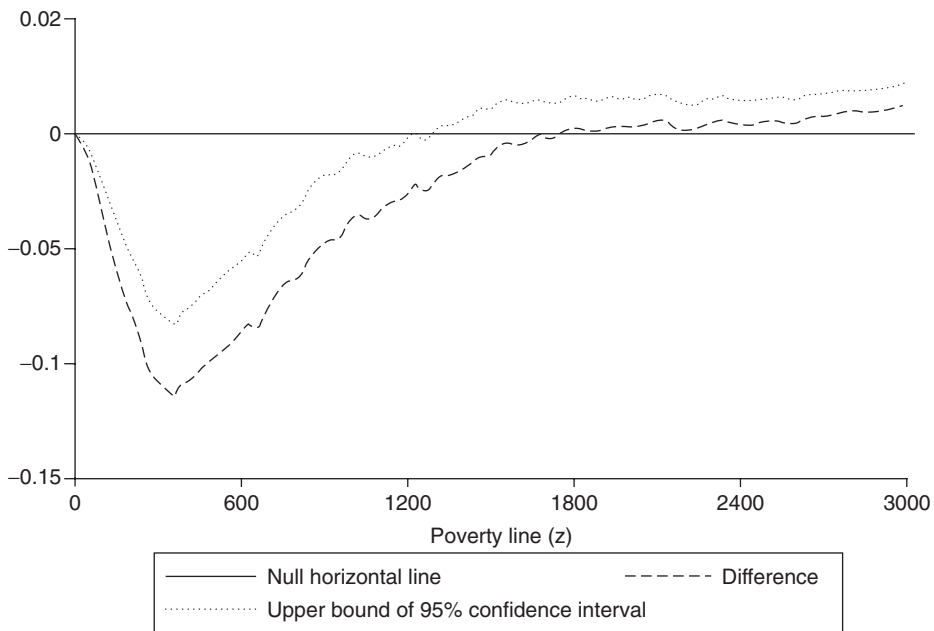
Figure 15. 1998 to 2004 is First-Order Relatively Pro-Poor. Difference
$P_{2004}((1 + g)z; \alpha = 0) - P_{1998}(z; \alpha = 0)$ is shown on the vertical axis

Figure 16. 1998 to 2004 is First-Order Relatively Pro-Poor. Difference $C_{2004}(p)/C_{1998}(p) - \mu_{2004}/\mu_{1998}$ is shown on the vertical axis

Figure 17. 1998 to 2004 is Strongly Second-Order Relatively Pro-Poor. Difference
$P_{2004}((1+g)z; \alpha = 1) - P_{1998}(z; \alpha = 1)$ is shown on the vertical axis

Figure 18. 1998 to 2004 is Strongly Second-Order Relatively Pro-Poor. Difference
$C_{2004}(p)/C_{1998}(p) - \mu_{2004}/\mu_{1998}$ is shown on the vertical axis

Figure 19. 1992 to 2004 is First-Order Relatively Pro-Poor. Difference $P_{2004}((1 + g)z; \alpha = 0) - P_{1992}(z; \alpha = 0)$ is shown on the vertical axis
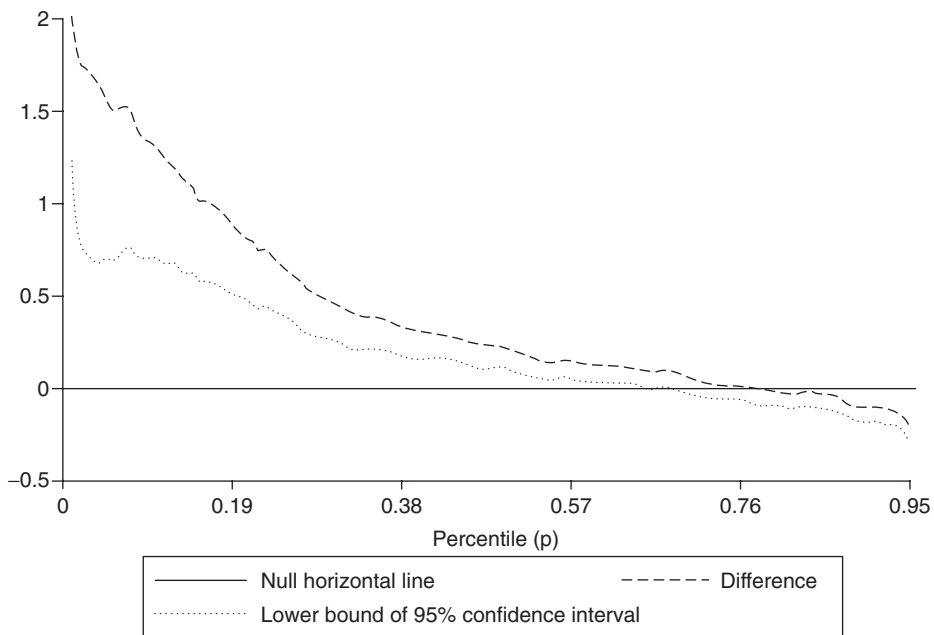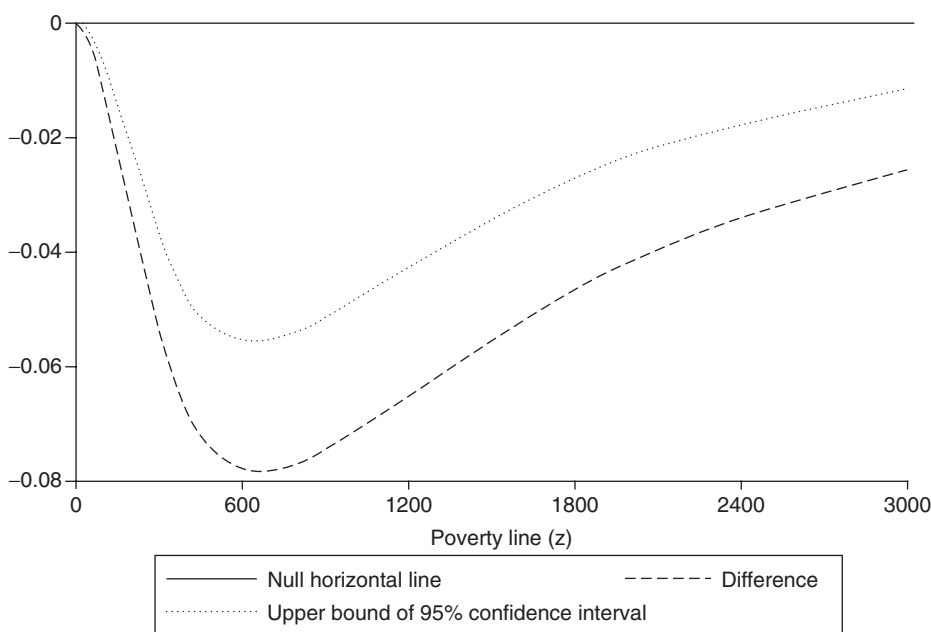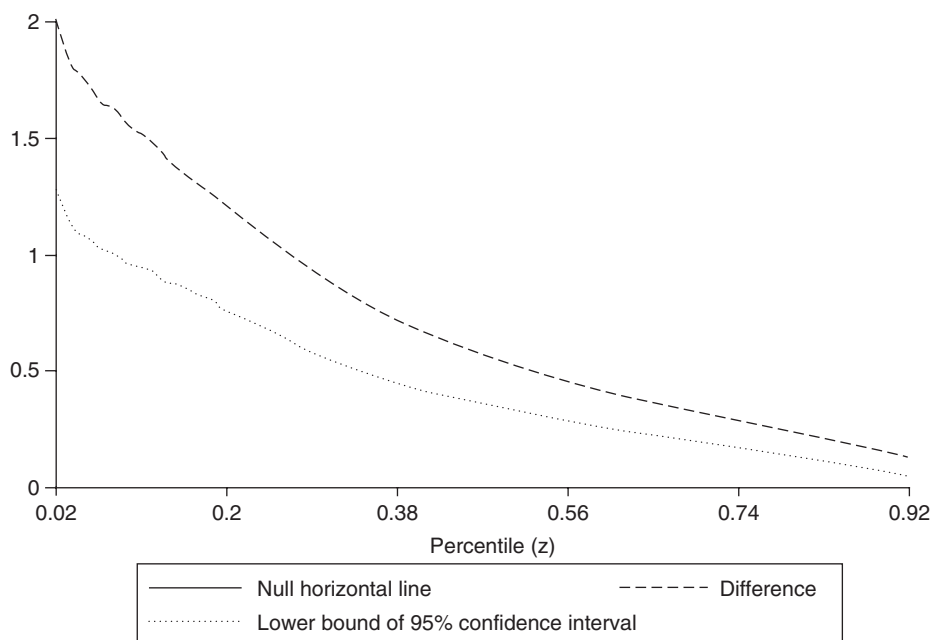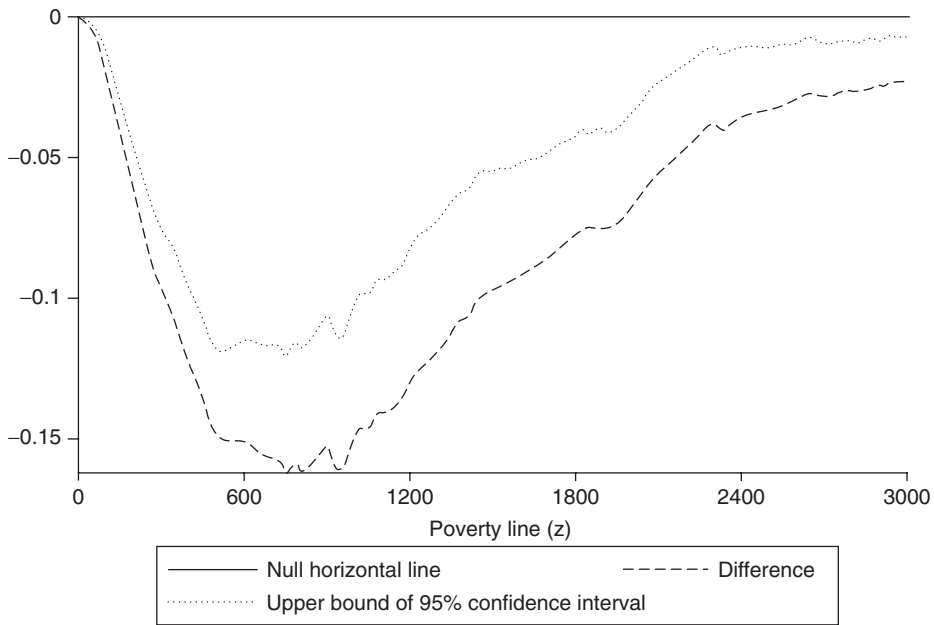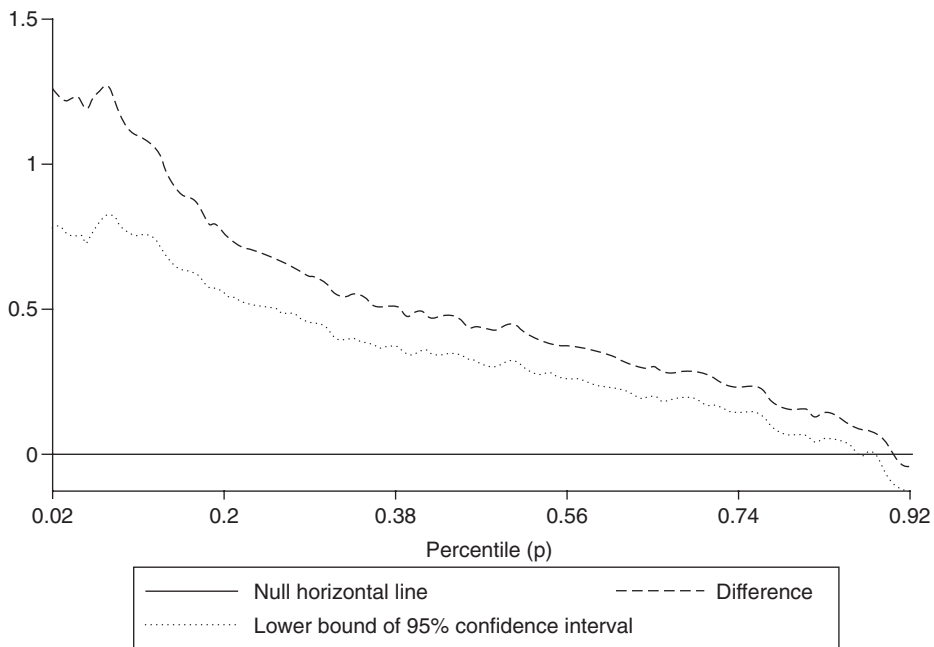

Figure 20. 1992 to 2004 is First-Order Relatively Pro-Poor. Difference $C_{2004}(p)/C_{1992}(p) - \mu_{2004}/\mu_{1992}$ is shown on the vertical axis

An analogous result is obtained in Figure 12 from comparing growth in quantiles to growth in average income. Again, for a substantial range of percentiles, the one-sided confidence interval overlaps with the zero line. As a referee has noted, tests at poverty lines or percentiles close to 0 can also be contaminated by relatively higher measurement errors in the lower tail of the distribution, especially if the extent of these measurement errors varies across time.

Moving to second-order relative pro-poorness does not help, as shown by Figures 13 and 14. The statistical insignificance now extends over a wider range or z (up to around 900 pesos) and p (up to around 0.4) values. This may seem counter-intuitive at first sight, but it follows from the fact that statistical uncertainty for first-order comparisons at the bottom of the distributions is compounded at the second-order since second-order conditions are made of cumulatives of first-order statistics (as discussed above). There is therefore an important lesson to be drawn here. If one were to omit testing for statistical significance, it might seem that second-order relative pro-poorness over the 1992–1998 period certainly cannot be weaker than first-order relative pro-poorness over the same period. But if one takes into account the effect of sampling variability at the bottom of the distribution, then the evidence for second-order relative pro-poorness is statistically weaker than that for first-order relative pro-poorness.

Testing for relative pro-poorness between 1998 and 2004 is more conclusive, as shown in Figures 15 and 16. The confidence interval around the sample estimates of $P_{2004}((1 + g)z; \alpha = 0) - P_{1998}(z; \alpha = 0)$ in Figure 15 is always below zero for z up to around 1200 pesos (as opposed to 1800 pesos for the sample estimates), which leads us to infer a robust first-order relative pro-poorness change in that period. A similar result is obtained in Figure 16 from comparing growth in quantiles to growth in average income. For a range of percentiles up to about 0.7, the lower bound of the confidence interval lies above the zero line.

Given the above results, it would seem interesting to test for second-order relative pro-poorness for the 1998–2004 period. The results are shown in Figures 17 and 18. We now obtain even stronger (and very strong) evidence of the relative pro-poorness of that period. The confidence interval is always below zero for differences $P_{2004}((1 + g)z; \alpha = 1) - P_{1998}(z; \alpha = 1)$ and above zero for differences $C_{2004}(p)/C_{1998}(p) - \mu_{2004}/\mu_{1998}$. This is not surprising given that, as discussed above, if first-order pro-poorness is verified statistically at order 1, then we can expect second-order pro-poorness also to be inferred statistically.

The results of the tests for relative pro-poorness over the period 1992 to 2004 are even stronger. These are shown in Figures 19 and 20. The confidence interval around the sample estimates of $P_{2004}((1 + g)z; \alpha = 0) - P_{1998}(z; \alpha = 0)$ in Figure 19 is always below zero even as we extend z beyond 3000 pesos. The same strong evidence is displayed in Figure 16 from the comparison of growth in quantiles to growth in average income between 1992 and 2004. For a range of percentiles up to about 0.9, the lower bound of the confidence interval is everywhere above the zero line. Hence, the period 1992–2004 shows a statistically significant and normatively solid degree of relative pro-poorness change in that period.

5. Conclusion

This paper proposes techniques to check for whether growth has been pro-poor. It first reviews different definitions of pro-poorness and argues for the use of methods that can generate results that are robust over classes of pro-poor measures and ranges of poverty lines. It then makes it empirically feasible to test for pro-poorness of growth. To do this, it derives the sampling distribution of the various estimators that are needed to test for absolute and relative pro-poorness. This leads to the convenient use of confidence intervals around the curves that must be ranked in order to conclude that a change has been robustly pro-poor—or anti-poor.

In doing this, the paper succeeds in implementing tests for the pro-poorness of growth (or recessions) that get around the difficulty of having to choose a poverty line to separate the poor from the non-poor, and of having to choose a set of normative weights to differentiate among the poor. It also provides novel tools for statistical inference in a robust normative framework.

The paper's techniques are then implemented using Mexico's National Income and Expenditure Surveys collected in 1992, 1998, and 2004, taking full account of the sampling design of these surveys and of the non-linearity and dependence of the estimators needed to test for pro-poorness. We find strong evidence that Mexican growth has been absolutely anti-poor between 1992 and 1998, absolutely pro-poor between 1998 and 2004 and between 1992 and 2004, and relatively pro-poor between 1992 and 2004 and between 1998 and 2004. This is useful and robust evidence on Mexico's distributional change during a period of significant economic turbulence and transformations.

The paper also demonstrates how and why observing pro-poorness in samples does not imply that we can immediately infer it for populations of interest (in contrast to earlier contributions to the pro-poorness literature). In particular, we find that the assessment of the 1992–1998 period is statistically too weak to lead to a robust pro-poor evaluation of that period. This is true for both first and second-order assessments of pro-poorness. An additional methodological finding is that, unlike what measurement theory implies for poverty dominance relationships *across populations*, moving to a higher order of dominance may lead to a weaker ability to order distributions statistically, when inference is based on the use of *samples*.

Appendix

A.1. *Calculation of cov* $(\hat{A}, \hat{B})$

To see how to compute var $(\hat{\Delta}^s(z) - \Delta^s(z))$ in (31) and (32), note from (29) that we can write

$$(40) \qquad \hat{A} = N_2^{-1} \sum_{h=1}^{N_2} \left(\frac{(1+g)z - y_2^h}{(1+g)z} \right)_+^\alpha - P_2((1+g)z; \alpha)$$

$$(41) \qquad + ((1+g)z) P_2^{gz}((1+g)z; \alpha)(\mu_2)^{-1} \left(N_2^{-1} \sum_{h=1}^{N_2} y_2^h - \mu_2 \right).$$

A similar expression can be obtained for $\hat{B}$ from (30):

$$(42) \qquad \hat{B} = N_1^{-1} \sum_{h=1}^{N_1} \left(\frac{z - y_1^h}{z} \right)_+^\alpha - P_1(z; \alpha)$$

$$(43) \qquad + ((1+g)z) P_2^{gz}((1+g)z; \alpha)(\mu_1)^{-1} \left(N_1^{-1} \sum_{h=1}^{N_1} y_1^h - \mu_1 \right).$$

Given the existence of the appropriate population moments of order 2, the estimators in (41) and (43) are root-N consistent by the central limit theorem. They are also asymptotically normal, with asymptotic covariance matrix given by:

$$
\begin{aligned}
(44) \quad \lim_{N \to \infty} N \operatorname{cov}\left(\hat{A}, \hat{B} \right) &= E\left[\left(\left(\frac{(1+g)z - y_2}{(1+g)z} \right)_+^\alpha + ((1+g)z) P_2^{gz}((1+g)z; \alpha)(\mu_2)^{-1} y_2 \right) \right. \\
&\quad \left. \cdot \left(\left(\frac{z - y_1}{z} \right)_+^\alpha + ((1+g)z) P_2^{gz}((1+g)z; \alpha)(\mu_1)^{-1} y_1 \right) \right] \\
&\quad - E\left[\left(\left(\frac{(1+g)z - y_2}{(1+g)z} \right)_+^\alpha + ((1+g)z) P_2^{gz}((1+g)z; \alpha)(\mu_2)^{-1} y_2 \right) \right] \\
&\quad \cdot E\left[\left(\left(\frac{z - y_1}{z} \right)_+^\alpha + ((1+g)z) P_2^{gz}((1+g)z; \alpha)(\mu_1)^{-1} y_1 \right) \right].
\end{aligned}
$$

An estimate of this expression can be computed using the sample observations. Special cases of cov $(\hat{A}, \hat{B})$ give var $(\hat{A})$ and var $(\hat{B})$, namely

$$
\begin{aligned}
(45) \quad \lim_{N \to \infty} N \operatorname{var}\left(\hat{A} \right) &= E\left[\left(\left(\frac{(1+g)z - y_2}{(1+g)z} \right)_+^\alpha + ((1+g)z) P_2^{gz}((1+g)z; \alpha)(\mu_2)^{-1} y_2 \right)^2 \right] \\
&\quad - E\left[\left(\left(\frac{(1+g)z - y_2}{(1+g)z} \right)_+^\alpha + ((1+g)z) P_2^{gz}((1+g)z; \alpha)(\mu_2)^{-1} y_2 \right) \right]^2
\end{aligned}
$$

and

$$
\begin{aligned}
(46) \quad \lim_{N \to \infty} N \operatorname{var}\left(\hat{B} \right) &= E\left[\left(\left(\frac{z - y_1}{z} \right)_+^\alpha + ((1+g)z) P_2^{gz}((1+g)z; \alpha)(\mu_1)^{-1} y_1 \right)^2 \right] \\
&\quad - E\left[\left(\left(\frac{z - y_1}{z} \right)_+^\alpha + ((1+g)z) P_2^{gz}((1+g)z; \alpha)(\mu_1)^{-1} y_1 \right) \right]^2.
\end{aligned}
$$

A.2. *Taking into Account Sampling Design*

The surveys we use in this paper (as well as most other real-world surveys) are not, however, drawn from the process of simple random sampling assumed above. Instead, they are stratified and clustered. Thus, let $s = 1, \ldots, L$ be the list of

the survey strata and n_s be the number of selected primary sampling units (PSU) in a stratum s. Further, let y_{si} be the sum of the observations of y in PSU si, $\overline{y_s} = n_s^{-1} \sum_{i=1}^{n_s} y_{si}$ be the mean of such y in stratum s, and let x_{si} and $\overline{x_s}$ be defined analogously for a variable x. Supposing that the number of PSU increases asymptotically to infinity, we can then estimate the sampling distribution of the above estimators $\hat{\Delta}^s(z)$ and $\hat{\Delta}^s(p)$ taking full account of the survey design. This can be done using the procedure described in Duclos and Araar (2006, pp. 284–7), a procedure that takes into account the sampling weights, the sampling design, and the number of statistical units (individuals) within each of the last sampling units (each household observation in the sample). Taking into account sampling design, the sampling covariance of two totals over the entire sample, $\hat{Y}$ and $\hat{X}$, is then estimated by

$$(47) \qquad \widehat{\text{cov}}_{SD}(\hat{Y}, \hat{X}) = \sum_{s=1}^{L} n_s^2 \frac{S_s^{yx}}{n_s}$$

where

$$(48) \qquad S_s^{yx} = (n_s - 1)^{-1} \sum_{i=1}^{n_s} \left(y_{si} - \overline{y_s} \right)\left(x_{si} - \overline{x_s} \right).$$

For the purposes of this paper, the above totals take the form of expressions such as $N_j \hat{\mu}_j$ in (21), $N_j \hat{P}_j((1+g)z; \alpha)$ in (24), and $N(\hat{Q}(p) - Q(p))$ in (35). The use of these expressions in (47) and (48) yields the sampling variances of the estimators $\hat{\mu}$, $\hat{\Delta}^s(z)$, $\hat{Q}(p)$ and $\hat{C}(p)$ in (21), (28), (35), and (36).

References

Alesina, A. and D. Rodrik, "Distributive Policies and Economic Growth," *Quarterly Journal of Economics*, 109, 465–90, 1994.

Atkinson, A., "On the Measurement of Inequality," *Journal of Economic Theory*, 2, 244–63, 1970.

Atkinson, A., B. Cantillon, E. Marlier, and B. Nolan, *Social Indicators—The EU and Social Inclusion*, Oxford University Press, Oxford, 2002.

Bahadur, R., "A Note on Quantiles in Large Samples," *Annals of Mathematical Statistics*, 37, 577–80, 1966.

Bourguignon, F., "The Poverty-Growth-Inequality Triangle," in *Conference on Poverty, Inequality and Growth*, Agence française de développement, Paris, 2003.

Bruno, M., M. Ravallion, and L. Squire, "Equity and Growth in Developing Countries: Old and New Perspectives on the Policy Issues," in Vito Tanzi and Ke-Young Chu (eds), *Income Distribution and High-Quality Growth*, MIT Press, Cambridge, MA, 1998.

Chakravarty, S., "Ethically Flexible Measures of Poverty," *Canadian Journal of Economics*, XVI, 74–85, 1983.

Clark, S., R. Hemming, and D. Ulph, "On Indices for the Measurement of Poverty," *The Economic Journal*, 91, 515–26, 1981.

Davidson, R. and J.-Y. Duclos, "Statistical Inference for the Measurement of the Incidence of Taxes and Transfers," *Econometrica*, 65, 1453–65, 1997.

Deininger, K. and L. Squire, "New Ways of Looking at Old Issues: Inequality and Growth," *Journal of Development Economics*, 38, 259–87, 1998.

Demombynes, G. and J. Hoogeveen, "Growth, Inequality and Simulated Poverty Paths for Tanzania, 1992–2002," *Journal of African Economies*, 16, 596–628, 2007.

Dollar, D. and A. Kraay, "Growth is Good for the Poor," *Journal of Economic Growth*, 7, 195–225, 2002.

Duclos, J.-Y., "What is Pro-Poor?" *Social Choice and Welfare*, 32, 37–58, 2009.

Duclos, J.-Y. and A. Araar, *Poverty and Equity Measurement, Policy, and Estimation with DAD*, Springer and IDRC, Berlin and Ottawa, 2006.

Duclos, J.-Y., P. Makdissi, and A. Araar, "Pro-Poor Tax Reforms, with an Application to Mexico," Mimeo, Université Laval, 2009.

Eastwood, R. and M. Lipton, "Pro-Poor Growth and Pro-Growth Poverty Reduction: What do they Mean? What does the Evidence Mean? What can Policymakers do?" *Asian Development Review*, 19, 1–37, 2001.

Essama-Nssah, B., "A Unified Framework for Pro-Poor Growth Analysis," *Economics Letters*, 89, 216–21, 2005.

Essama-Nssah, B. and P. J. Lambert, "Measuring Pro-Poorness: A Unifying Approach with New Results," *Review of Income and Wealth*, 55, 752–78, 2009.

Foster, J. and A. Shorrocks, "Inequality and Poverty Orderings," *European Economic Review*, 32, 654–61, 1988a.

———, "Poverty Orderings," *Econometrica*, 56, 173–7, 1988b.

Foster, J., J. Greer, and E. Thorbecke, "A Class of Decomposable Poverty Measures," *Econometrica*, 52, 761–76, 1984.

Kakwani, N. and E. Pernia, "What is Pro Poor Growth?" *Asian Development Review*, 18, 1–16, 2000.

Kakwani, N. and H. H. Son, "Poverty Equivalent Growth Rate," *Review of Income and Wealth*, 54, 643–55, 2008.

Kakwani, N., S. Khandker, and H. Son, "Poverty Equivalent Growth Rate: With Applications to Korea and Thailand," Technical Report Economic Commission for Africa, 2003.

Klasen, S., "In Search of the Holy Grail: How to Achieve Pro-Poor Growth?" in B. Tungodden, N. Stern, and I. Kolstad (eds), *Toward Pro Poor Policies-Aid, Institutions, and Globalization*, Oxford University Press, New York, 63–94, 2004.

———, "Economic Growth and Poverty Reduction: Measurement Issues in Income and Non-Income Dimensions," *World Development*, 36, 420–45, 2008.

McCulloch, N. and B. Baulch, "Tracking Pro-Poor Growth," Technical Report ID21 Insights #31, Institute of Development Studies, Sussex, 1999.

Nicita, A., "Who Benefited from Trade Liberalization in Mexico? Measuring the Effects on Household Welfare," Policy Research Working Paper Series 3265, The World Bank, 2004.

Rao, R., *Linear Statistical Inference and Its Applications*, John Wiley and Sons, New York, 1973.

Ravallion, M. and S. Chen, "Measuring Pro-poor Growth," *Economics Letters*, 78, 93–9, 2003.

Ravallion, M. and G. Datt, "Why Has Economic Growth Been More Pro-poor in Some States of India Than Others?" *Journal of Development Economics*, 68, 381–400, 2002.

Rawls, J., *A Theory of Justice*, Harvard University Press, Cambridge, MA, 1971.

Shorrocks, A., "Ranking Income Distributions," *Economica*, 50, 3–17, 1983.

Son, H., "A Note on Pro-Poor Growth," *Economics Letters*, 82, 307–14, 2004.

Szekely, M., "Pobreza, desigualdad y entorno macroeconomico en Mexico," *Informacion Comercial Española Revista de Economia*, 125–42, 2005a.

———, "Pobreza y desigualdad en Mexico entre 1950 y 2004," *El Trimestre Economico*, 72, 913–31, 2005b.

Tétreault, D., "The Evolution of Poverty in Late 20th-Century Mexico," *Canadian Journal of Development Studies*, 27, 309–26, 2006.

United Nations, *A Better World for All*, United Nations, New York, 2000.

Watts, H. W., "An Economic Definition of Poverty," in D. Moynihan (ed.), *Understanding Poverty*, Basic Books, New York, 1968.

World Bank, *The Quality of Growth*, Oxford University Press, New York, 2000.