Socially-Efficient Tax Reforms*

Jean-Yves Duclos[†], Paul Makdissi[‡] and Quentin Wodon[§]

October 2001

Abstract

This paper proposes graphical methods to determine whether commodity-tax changes are "socially efficient", in the sense of improving social welfare or decreasing poverty for large classes of social welfare and poverty indices. It also derives estimators of critical poverty lines and economic efficiency ratios which can be used to characterize socially-efficient tax reforms. The statistical properties of the various estimators are derived in order to make the method implementable using survey data. The methodology is illustrated using a recently-proposed reform of the Mexican Valued Added Tax system.

Keywords: Social welfare, Poverty, Efficiency, Tax Reform, Stochastic Dominance
JEL Numbers: D12, D63, H21, I32

* We thank R. Alessie, J.O. Lanjouw, P. Lanjouw, M. Ravallion, S. Yitzhaki for many useful comments. This paper was funded through the World Bank Research Support Budget under the research project "The impact of changes in prices, taxes, subsidies and stipends on poverty when household differ in needs" and has also benefitted from the support of SSHRC, the MIMAP programme of IDRC and Bureau de la recherche of Université de Sherbrooke.

† CRÉFA, Université Laval, Québec, and Institut d'Anàlisi Econòmica, CSIC, Barcelona. Corresponding address: CRÉFA, Département d'économique, Pavillon De Sève, Université Laval, Sainte-Foy, Québec, Canada, G1K 7P4. Email: jyves@ecn.ulaval.ca

‡ Département d'économique and CEREF, Université de Sherbrooke, 2550 boulevard de l'Université, Sherbrooke, Québec, Canada, J1K 2R1; email: paul.makdissi@courrier.usherb.ca

§ LCSPR, World Bank, 1818 H Street, NW, Washington, DC 20433, USA, Email: qwodon@worldbank.org.

1 Introduction

Policies that affect consumer and producer prices have an impact on welfare, and there are many such policies. Some governments maintain high import tariffs or fail to implement regulations to foster competition. This may protect national producers by maintaining high domestic producer prices, but it also raises consumption prices, which hurts consumers. Most governments use sales and indirect taxes to raise tax revenues, thus affecting consumer and producer prices. This is particularly important for developing countries, which rely heavily on commodity taxation to generate tax revenues. Price subsidies on food, education, energy and transportation are also common in developing and developed countries alike.

Policy analysts must routinely evaluate the impact of such pricing policies on poverty and social welfare. Important information problems stand, however, in the way of anyone wishing to do this. The main objective of the paper is to show how these problems can be (somewhat) circumscribed. We are particularly interested in demonstrating the empirical applicability of a "social efficiency" approach, which makes it possible to identify price changes that will be deemed socially desirable by wide spectra of social welfare and poverty analysts. For expositional simplicity, we will focus in this paper on the analysis of the effect of indirect tax reforms on social welfare and poverty. We will think of two goods, j and l, and will ask whether it is socially desirable (in a sense to be defined precisely below) to increase a tax rate on good j in order to decrease a tax rate on good l. Note that tax rates may be positive or negative, which also says that decreasing the tax on a good may mean increasing its rate of subsidy. In doing this, we keep producer prices constant for expositional simplicity, but this framework can be extended to the general equilibrium analysis of any government intervention that changes relative prices.

The first measurement difficulty is then to estimate the impact of price changes on consumer welfare. It is well known that this can be a complicated exercise and that its results can be sensitive to a number of important theoretical and econometric assumptions. The task is particularly problematic when the goal is to find globally optimal tax systems, since the tax analyst then needs reliable demand responses over extended price intervals for each household. Keeping in mind that actual changes in the tax system are "slow and piecemeal" (see for instance Feldstein (1975)), and that it would be unwise to ignore the role of the actual tax system as a departure (or anchor) point for the identification of desirable tax reforms, we focus here on the effect of marginal tax reforms. One immediate advantage is that evaluating the distributive impact of marginal tax reforms does not require estimates of individual demand and utility functions, but can instead be assessed directly from the observed data alone[1], as we review in Section 2.

The second difficulty resides in the choice of a social evaluation function with which the impact of the tax reform is to be measured[2]. This choice poses a fundamental problem, since any particular selection of functional form and parameters for a social evaluation function necessarily embodies arbitrary value judgements. The strategy we follow in this paper is instead to define classes of social evaluation indices that incorporate increasingly stronger judgements on the importance of distributive issues in designing tax policy. For this purpose, we consider social evaluation indices that take into account (though not necessarily with the same distributive weight) everyone's welfare – the traditional social welfare functions – as well as social evaluation indices that censor welfare at a poverty line and which are therefore not affected by changes in the welfare of the "rich" – in short, the much-used poverty indices. These classes of functions and indices are described in Section 3.

Verifying whether a tax reform is desirable then proceeds by checking whether it can command the unanimous approval of all those analysts who agree with some generally-defined normative properties of the social evaluation indices. If so, the reform can be called "socially efficient". One well-known criterion for social efficiency property comes from the Pareto principle. Pareto efficiency, however, is likely to provide a poor normative basis for the empirical analysis of tax reforms (for reasons that we discuss below in Section 7), and it is thus useful to consider "higher-order" ethical principles. We increase progressively the ethical content of our classes of social welfare and poverty indices by incorporating the anonymity – or symmetry – principle (for first-order efficiency), the Pigou-Dalton principle of transfers (for second-order efficiency), the principle of diminishing transfers (for third-order

[1] The analysis of marginal tax reforms has also been considered among others by Guesnerie (1977), Ahmad and Stern (1984, 1991), Besley and Kanbur (1988), Yitzhaki and Thirsk (1990), Yitzhaki and Slemrod (1991), Mayshar and Yitzhaki (1995), Yitzhaki and Lewis (1996) and Makdissi and Wodon (2002).

[2] See Deaton (1977), King (1983) and Besley and Kanbur (1988) for examples of the use of particular social evaluation functions in assessing tax reforms, and Christiansen and Jansen (1978) and Ahmad and Stern (1984) for the estimation of "implicit social preferences".

efficiency), and subsequent "generalized" higher-order normative principles. These principles lead to the definition of what we term Pen-efficient, Dalton-efficient, Kolm-efficient and higher-order welfare-efficient tax reforms. The definition of poverty-efficient tax reforms proceeds in a similar way by censoring the social assessment at a poverty line. This is derived in Section 4.

The third difficulty lies in estimating the impact on tax revenues of changes in the structure of indirect taxation. It is well known that this impact is linked to the aggregate deadweight loss of taxation (Wildasin (1984) and Mayshar (1990)), and thus to the economic efficiency of a tax reform. Estimating the aggregate impact on tax revenues can be less difficult (in principle at least) than the estimation of household-specific changes in consumption, but the details of the estimation procedures can still lead to considerable disagreements among tax analysts. One solution is to carry out sensitivity analysis on the role of the unobserved economic efficiency parameter, in the manner of Ahmad and Stern (1984) for instance. We propose an alternative procedure that estimates the critical efficiency ratio up to which a tax reform can be said to be socially efficient at a given ethical order. This leaves policy makers free to assess whether the actual efficiency ratio is likely (or can be safely estimated) to be below that critical value, and thus whether the tax reform can confidently be deemed socially efficient. A similar device is constructed to handle, for poverty measurement, the role of poverty lines – whose estimation is also notoriously difficult and controversial. We show how to construct estimates of the critical poverty lines up to which a tax reform can be considered to be poverty efficient. Policy makers and poverty analysts are then let free to assess or estimate whether those critical poverty lines are sufficiently high to encompass all plausible poverty line estimates. If that is so, the tax reform can be confidently described as poverty efficient. These critical thresholds are derived in Section 5.

Social efficiency is checked in this paper through the use of simple Consumption Dominance ($\overline{CD}$) curves. $\overline{CD}$ curves display cumulative consumption shares when these are weighted by powers of poverty gaps. First-order $\overline{CD}$ curves show the share in the total consumption of a good of those at a given income level. Second-order $\overline{CD}$ curves indicate the combined share in the total consumption of a good of those whose income lies below a given threshold. Higher-order $\overline{CD}$ curves weight consumption shares by increasingly higher powers of poverty gaps. Increasing the tax on good j and decreasing the tax on good l is poverty efficient at any given order of ethical dominance if the $\overline{CD}$ curve of that order for good l is higher than the $\overline{CD}$ curve for good j at every threshold under a maximum poverty line. When that maximum poverty line extends to infinity, the tax reform can be called welfare efficient. The second-order welfare efficiency condition we obtain is equivalent to the Dalton efficiency condition of Yitzhaki and Slemrod (1991) and Mayshar and Yitzhaki (1995). Our social efficiency conditions become less stringent when we consider poverty instead of welfare efficiency, since we can then focus exclusively on those below a maximum admissible poverty line and do not need to consider the impact of the reform on the entire population. Increasing the order of dominance also facilitates the search for socially efficient tax reforms since more ethical structure is then imposed on the properties of the admissible social evaluation indices. These links are described in some detail in Section 6.

$\overline{CD}$ curves were first used by Makdissi and Wodon (2002) in the context of poverty reduction. This paper uses them for the broader analysis of social efficiency. The current paper also introduces the concept of first-order social efficiency and shows how to test it, and also discusses how it differs from the traditional concept of Pareto efficiency. The concept of social efficiency further leads to an interpretation of the ratio of $\overline{CD}$ curves as a natural indicator of the distributive benefit of a tax reform, a benefit which can be compared to the deadweight loss or economic efficiency cost of the reform. We also see how the concepts of critical economic efficiency ratios and critical poverty lines may help circumvent the difficulties faced by tax analysts.

Tax policy analysis is clearly of practical and policy interest, and it is thus essential to consider how the tools described above can be implemented empirically using survey data. We do this in part in this paper by investigating in Section 8 the sampling properties of estimators of $\overline{CD}$ curves, critical poverty lines and critical economic efficiency thresholds. In this, we make use *inter alia* of non-parametric regression curves for the analysis of first-order-efficient tax reforms.

The paper's tools are applied in Section 9 to the analysis of a recent proposal for the reform of the Mexican Value-Added-Tax system. Section 10 concludes the paper, and the proofs of the various lemmas and theorems can be found in the Appendix.

2 Notation and definitions

Consider a vector q of K consumer prices. For expositional simplicity, and as is customary in the partial equilibrium literature, we set the vector of producer prices to 1 and assume them to be constant and invariant to changes in t, the vector of K tax rates. We then have $q = 1 + t$ and $dq_k = dt_k$, where q_k and t_k denote respectively the price of and the tax rate on good k.

Let y be exogenous income, and denote consumers' preferences by θ. The indirect utility function is given by $v(y, \theta, q)$. Following King (1983), we will be using a vector of reference prices, q^R, to assess consumers' well-being in the presence of varying tax rates. Denote the real (or equivalent) income in the post-reform situation by y^R, where y^R is measured on the basis of the reference prices q^R. y^R is implicitly defined by $v\left(y^R, \theta, q^R\right) = v\left(y, \theta, q\right)$, and explicitly by the real income function $y^R = \rho\left(y, \theta, q, q^R\right)$, where

$$v\left(\rho\left(y, \theta, q, q^R\right), \theta, q^R\right) \equiv v\left(y, \theta, q\right). \tag{1}$$

By definition, y^R gives the level of income that provides under q^R the same utility as y yields under q. The nominal income function, $y = \eta\left(y^R, \theta, q, q^R\right)$, is the inverse of $\rho\left(y, \theta, q, q^R\right)$ and is such that

$$v\left(y^R, \theta, q^R\right) \equiv v\left(\eta\left(y^R, \theta, q, q^R\right), \theta, q\right). \tag{2}$$

The nominal income function gives the level of income that yields the same utility under q as y^R gives under q^R.

We then wish to determine how consumer welfare is affected by a marginal change in tax rates. Let $x_k\left(y, \theta, q\right)$ be the consumption of good k of a consumer with income y, preferences θ and facing prices q. Using Roy's identity and setting reference prices to pre-reform prices, we find:

$$\begin{aligned}
\left.\frac{\partial \rho(y, \theta, q, q^R)}{\partial t_k}\right|_{q=q^R} &= \left.\frac{\partial v\left(y, \theta, q\right)/\partial t_k}{\partial v\left(y^R, \theta, q^R\right)/\partial y^R}\right|_{q=q^R} \\
&= -\left.\frac{\partial v\left(y, \theta, q\right)/\partial y}{\partial v\left(y^R, \theta, q^R\right)/\partial y^R}\right|_{q=q^R} \cdot x_k\left(y, \theta, q\right) \\
&= -x_k\left(y, \theta, q^R\right).
\end{aligned} \tag{3}$$

Similarly,

$$\left.\frac{\partial \eta\left(y^R, \theta, q, q^R\right)}{\partial t_k}\right|_{q=q^R} = -x_k\left(y, \theta, q^R\right). \tag{4}$$

Equations (3) and (4) say that observed pre-reform consumption of good k is a sufficient statistic to know the impact on consumer welfare of a marginal change in the price of good k [3].

Assume that preferences θ and exogenous income y are jointly distributed according to the distribution function $F(y,\theta)$. The conditional distribution of θ given y is denoted by $F(\theta\,|y)$, and the marginal distribution of nominal income is given by $F(y)$. Continuity across y of the conditional distribution function $F(\theta\,|y)$ is not generally needed for the methodological results of this paper, but empirical applicability of some of these results will sometimes require such a continuity assumption, as we will discuss explicitly in Section 8. Let preferences belong to the set Ω, and assume income to be distributed over $[0,a]$. Expected consumption of good k at income y is given by $x_k(y,q)$, such that

$$x_k(y,q) = \underset{\theta}{E}\left[x_k\left(y,\theta,q\right)\right] = \int_{\Omega} x_k\left(y,\theta,q\right)\, dF(\theta|y). \tag{5}$$

Let $X_k(q)$ then be the *per capita* consumption of the k^{th} good, defined as $X_k(q) = \int_0^a x_k\left(y,q\right)\, dF\left(y\right)$. By (3) and (4), $X_k(q)$ is also the average welfare cost of an increase in the price of good k. We will often normalize by $X_k(q)$ some of the various measures that will be associated to a good k, and we will distinguish these normalized measures with a $\bar{\cdot}$. Hence, as a proportion of *per capita* consumption, consumption of good k at income y is expressed as $\overline{x}_k(y,q) = x_k(y,q)/X_k(q)$.

We now turn to the government budget constraint. *Per capita* commodity tax revenues, $R(q)$, equal $R(q) = \sum_{k=1}^{K} t_k X_k(q)$. Without loss of generality, assume that the government's tax reform increases the tax rate on the j^{th} commodity and uses the proceeds to decrease the tax rate (or to increase the subsidy) on the l^{th} commodity. As is conventional in the optimal taxation literature, total tax revenues are kept invariant to the tax reform. Revenue neutrality of the tax reform requires that

$$dR(q) = \left[X_j(q) + \sum_{k=1}^{K} t_k \frac{\partial X_k(q)}{\partial q_j}\right] dq_j + \left[X_l(q) + \sum_{k=1}^{K} t_k \frac{\partial X_k(q)}{\partial q_l}\right] dq_l = 0. \tag{6}$$

[3] Note that (3) and (4) are valid for rationed goods too.

Define γ as

$$\gamma = \frac{X_l + \sum_{k=1}^{K} t_k \frac{\partial X_k}{\partial q_l}}{X_l} \bigg/ \frac{X_j + \sum_{k=1}^{K} t_k \frac{\partial X_k}{\partial q_j}}{X_j}\ , \tag{7}$$

where, for expositional simplicity, the (q)'s have been omitted. The numerator in (7) gives the marginal tax revenue of a marginal increase in the price of good l, per unit of the average welfare cost that this price increase imposes on consumers. Equivalently, this is 1 minus the deadweight loss of taxing good l, or the inverse of the marginal cost of public funds (MCPF) from taxing l (see Wildasin (1984)). The denominator gives exactly the same measures for an increase in the price of good j. γ is thus the economic (or "average") efficiency of taxing good l relative to taxing good j. We may thus interpret γ as the efficiency cost of taxing j relative to that of taxing l (the MCPF for j over that for l). The higher the value of γ, the less economically efficient is taxing good j.

By simple algebraic manipulation, we can then rewrite equation (6) as

$$dq_j = -\gamma \left(\frac{X_l}{X_j} \right) dq_l, \tag{8}$$

which fixes dq_j as a revenue-neutral proportion of dq_l.

3 Measuring poverty and social welfare

To assess the impact of a tax reform on poverty and social welfare, we follow the main custom in the measurement literature and focus for simplicity on additive poverty and social welfare indices. For first, second and third order dominance, this is only to ease the necessity and sufficiency proofs of the various social efficiency conditions: we could use some of the results of the previous literature to show that the relevant classes of indices also include subclasses of non-additive indices as well[4]. This additivity assumption is

[4] See, for instance, Dasgupta, Sen and Starrett (1973) and Foster and Shorrocks (1988a,b).

such poverty indices can be expressed as:

$$P(z) = \int_0^a p(y,z)\, dF(y) \tag{9}$$

where $P(z)$ is an additive poverty index, z is the poverty line defined in income space and $p(y,z)$ is the contribution to total poverty of a consumer with income y (with $p(y,z) = 0$ for all $y > z$)[5].

For the poverty efficiency results of this paper, we consider the classes of poverty indices $P(z) \in \Pi^s$ defined such that

$$\Pi^s(z) = \left\{ P(z) \;\middle|\; \begin{array}{c} p(y,z) = 0 \text{ if } y > z,\; p(y,z) \in \widehat{C}^s(z), \\ (-1)^i p^{(i)}(y,z) \geq 0 \text{ for } i = 0, 1, ..., s, \\ p^{(t)}(z,z) = 0 \text{ for } t = 0, 1, ..., s-2 \text{ when } s \geq 2 \end{array} \right\} \tag{10}$$

where $\widehat{C}^s(z)$ is the set of functions that are s-time piecewise differentiable[6] over $[0, z]$, and where the subscript $^{(s)}$ stands for an s^{th}-order derivative with respect to y. We will return very shortly to the interpretation of the derivative assumptions.

A particular subclass of additive poverty indices to which we will make repeated references below is found in Foster, Greer and Thorbecke (1984) and is defined for $\alpha \geq 0$ by

$$FGT^\alpha(z) = \int_0^z \left(\frac{z-y}{z} \right)^\alpha dF(y). \tag{11}$$

[5]It is best to think of y and z as *real* variables. Defining the poverty line in the real income space rather than in the nominal income space is convenient since the poverty line is then invariant to tax reforms. In this paper, as mentioned above, we generally suppose that pre-reform nominal and real incomes are the same since we take reference prices as the pre-reform prices (this is the common – though arbitrarily-made – assumption in the literature; see Donaldson (1992) for a general discussion). In any case, the marginal distribution of real income, $F_R(y^R)$, can always be computed from the joint distribution of nominal income and preferences $F(y, \theta)$:

$$F_R(y^R) = \int_\Omega \int_0^{\eta(y^R, \theta, Q)} dF(y, \theta).$$

[6]When the $(s-1)^{th}$ derivative is a piecewise differentiable function, the function and its $(s-2)$ first derivatives are differentiable everywhere. The $\widehat{C}^s(z)$ continuity assumption is made for analytical simplicity since it could be relaxed to include indices whose $(s-1)^{th}$ derivative is discontinuous (and which are therefore not s-time piecewise differentiable).

$FGT^0(z)$ gives the most widely used index of poverty, the so-called poverty headcount, and $FGT^1(z)$ yields the second most popular index, the (normalized) average poverty gap. Note that $FGT^\alpha(z)$ belongs to $\Pi^s(z)$ for $\alpha \geq s-1$. The FGT indices were in fact used by Besley and Kanbur (1988) for their analysis of price subsidies (thus in fact supposing an explicit form for $p(y,z)$). Other well-known additive indices that also belong to $\Pi^1(z)$ and $\Pi^2(z)$ include the Watts (1968) index, the second class of indices proposed by Clark, Hemming and Ulph (1981), and the class of indices proposed by Chakravarty (1983).[7]

Turning now to social welfare, we consider utilitarian social welfare functions U such that:

$$U = \int_0^a u(y)\, dF(y). \tag{12}$$

We focus on social welfare indices that belong to the classes Ω^s, $s = 1, 2, \ldots$. They are defined as

$$\Omega^s = \left\{ U \middle| u(y) \in C^s(\infty),\ (-1)^{i+1} u^{(i)}(y) \geq 0 \text{ for } i = 1, 2, ..., s, \right\}. \tag{13}$$

It is useful at this point to provide a normative interpretation of the different classes of indices of poverty and social welfare. When $s = 1$, poverty indices weakly decrease ($p^{(1)}(y,z) \leq 0$) while welfare indices weakly increase ($u^{(1)}(y) \geq 0$) when an individual's income increases. These indices are thus Paretian but they also obey the well-known symmetry or anonymity axiom: interchanging any two individuals' incomes leaves unchanged the poverty and social welfare indices. Ordering two distributions of living standards over the first-order classes of indices is equivalent to making the living standards "parade" simultaneously alongside each other, and verifying if one parade weakly dominates the other (this exercise was first suggested by Pen (1971, ch.III)). For poverty comparisons, the distributions of living standards are simply censored at z. For simplicity, we will refer to first-order welfare-efficient tax reforms as "Pen-efficient tax reforms".

When $s = 2$, poverty indices are convex and welfare indices are concave. They must thus respect the Pigou-Dalton principle of transfers, which postulates that a mean-preserving transfer of income from a higher-income

[7] As pointed out in Zheng (1999), it is possible to transform some of those indices in order to make them satisfy higher-order conditions. Relaxing the additivity or piecewise differentiability assumptions would also include other indices in the $\Pi^s(z)$ classes.

person to a lower-income person constitutes a social improvement, in the form of increasing social welfare or decreasing poverty. We follow the lead of Mayshar and Yitzhaki (1995) by denoting as "Dalton-efficient tax reforms" those reforms that will be found to be second-order welfare efficient.

By their third-order derivative, the poverty and social welfare indices that belong to Π^3 and Ω^3 must also be sensitive to favorable composite transfers. These transfers are such that a beneficial Pigou-Dalton transfer within the lower part of the distribution, accompanied by an adverse Pigou-Dalton transfer within the upper part of the distribution, will add to social welfare, provided that the variance of the distribution is not increased. Kolm (1976) was the first to introduce this condition into the inequality literature, and we therefore refer for simplicity to third-order welfare efficient tax reforms as "Kolm-efficient tax reforms". Kakwani (1980) subsequently adapted this principle to poverty measurement[8].

For the interpretation of higher orders of dominance, we can use the generalized transfer principles of Fishburn and Willig (1984). For $s = 4$, for instance, consider a combination of two exactly opposite and symmetric composite transfers, the first one being favorable and occurring within the lower part of the distribution, and the second one being unfavorable and occurring within the higher part of the distribution. Because the favorable composite transfer occurs lower down in the income distribution, indices that are members of the $s = 4$ classes will respond favorably to this combination of composite transfers. Generalized higher-order transfer principles essentially postulate that, as s increases, the weight assigned to the effect of transfers occurring at the bottom of the distribution also increases. Blackorby and Donaldson (1978) describe these indices as becoming more Rawlsian. Thus, as shown in Davidson and Duclos (2000) for poverty indices, when $s \to \infty$ only the lowest income counts, although this result does not generalize to welfare dominance, as shown in Duclos and Makdissi (2000) and as will also appear later.

[8]See also Shorrocks and Foster (1987) for a characterization of the composite transfer principle and Davies and Hoy (1994) for a description on the normative implications of this principle.

4 Identifying Socially Efficient Tax Reforms

To derive the conditions by which the social efficiency of tax reforms can be checked, it is handy first to refer to stochastic dominance curves. In real income space, these are defined as

$$D^s(z) = \frac{1}{(s-1)!} \int_{\Omega} \int_0^{\eta(z,\theta,Q)} \left[z - \rho\left(y,\theta,Q\right)\right]^{(s-1)} dF\left(y,\theta\right), \tag{14}$$

for orders of dominance $s = 1, 2, ...$ When $q = q^R$, we have that $z = \eta\left(z,\theta,q,q^R\right) = \rho\left(z,\theta,q,q^R\right)$, and (14) then reduces to the simpler expression

$$D^s(z) = \frac{1}{(s-1)!} \int_0^z \left[z - y\right]^{(s-1)} dF\left(y\right). \tag{15}$$

Dominance curves are therefore just sums of powers of poverty gaps. They can thus be interpreted as ethically-weighted sums of individual deprivation. The larger the value of s, the larger the weights on the largest poverty gaps. Clearly, as can be seen by comparing (11) and (15), dominance curves have a convenient link with the FGT indices since $FGT^{\alpha}(z) = (\alpha)! z^{-\alpha} D^{\alpha+1}(z)$.

It is thus useful to consider how the dominance curves are affected by changes in prices. By (3), (4) and (14), we can show that:

$$\left.\frac{\partial D^s\left(z\right)}{\partial t_k}\right|_{q=q^R} = \begin{cases} x_k\left(z,q^R\right)\, f(z), & \text{if } s = 1 \\ \frac{1}{(s-2)!}\int_0^z x_k\left(y,q^R\right)\,\left(z-y\right)^{s-2}\, dF\left(y\right) & \text{if } s = 2,3,... \end{cases} \tag{16}$$

where $f(z)$ is the density of income at z. For reasons that will become clear later, these derivatives can serve to define "consumption dominance" (CD) curves:

$$CD_k^s\left(z\right) = \frac{\partial D^s\left(z\right)}{\partial t_k}, s = 1, 2, ... \tag{17}$$

Because $CD_k^s\left(z\right)$ curves describe changes in ethically weighted sums of deprivation, they can be interpreted as the ethically weighted cost of taxing k. Normalized CD curves, $\overline{CD}_k^s(z)$, are just the above CD curves for good k

normalized by the average consumption of that good:

$$\overline{CD}_k^s(z) = \frac{CD_k^s(z)}{X_k(q)}. \tag{18}$$

$\overline{CD}$ curves are thus the ethically weighted (or social) cost of taxing k as a proportion of the average welfare cost. Note that the social cost depends on s and z. By (16), $\overline{CD}_k^1(z)$ only takes account of the consumption pattern of those precisely at z. The $\overline{CD}_k^2(z)$ curve gives the share in the total consumption of k of those individuals with income less than z. For $s = 3, 4, ...$, greater weight is given to the shares of those with higher poverty gaps. For $z < a$, the pattern of consumption and the welfare cost for some may be ignored in the computation of the social cost of taxing k. The social and average welfare cost coincide only when $s = 2$ and $z \geq a$, since we have $\overline{CD}_k^2(z) = 1$ for all $z \geq a$ and all k.

Define a distribution function for the consumption of good k as

$$G_k(y) = \int_0^y \overline{x}(u, q)\, f(u)\, du. \tag{19}$$

$G_k(y)$ is the proportion of the total consumption of good k that is consumed by those with incomes less than y. Note that $G_k(a) = 1$. For $s \geq 2$, it follows from (16) that

$$\overline{CD}_k^s(z) = \frac{1}{(s-2)!} \int_0^z (z-y)^{s-2}\, dG_k(y). \tag{20}$$

Hence, the $\overline{CD}^s$ curves are scaled FGT^α indices, using $\alpha = s - 2$ and a transformed income distribution G_k that weights individuals by their share in the total consumption of the good k. When multiplied by $(s-1)!/z^{s-1}$, the CD curves for $s = \alpha + 1$ have the convenient feature of providing the impact on the $FGT^\alpha(z)$ indices of a marginal increase in the price of good k. This is important in its own right given the popularity of the FGT class of poverty indices.

Most importantly it would seem, Consumption Dominance curves can be used to test the poverty and welfare efficiency of tax reforms. This is shown in Theorems 1 and 2, where γ is defined as in (7).

Theorem 1 *A necessary and sufficient condition for a marginal tax reform, $dq_j = -\gamma \left(\frac{X_l}{X_j}\right) dq_l > 0$, to be s-order poverty efficient, that is, to decrease poverty weakly for all $P(z) \in \Pi^s(z)$, for all $z \in [0, z^+]$ and for a given $s \in \{1, 2, 3, ...\}$, is that*

$$\overline{CD}_l^s(y) - \gamma \overline{CD}_j^s(y) \geq 0,\ \forall y \in \left[0, z^+\right]. \tag{21}$$

Proof. See appendix. ■

Theorem 1 is similar to the result obtained in Makdissi and Wodon (2000), except for the case of first-order poverty efficiency, which they did not consider. For welfare dominance, a similar theorem applies.

Theorem 2 *A sufficient condition for a marginal tax reform, $dq_j = -\gamma \left(\frac{X_l}{X_j}\right) dq_l > 0$, to be s-order welfare efficient, that is, to increase social welfare weakly for all $W \in \Omega^s$ and for a given $s \in \{1, 2, 3, ...\}$, is that*

$$\overline{CD}_l^s(y) - \gamma \overline{CD}_j^s(y) \geq 0,\ \forall y \in [0, \infty) \tag{22}$$

Proof. See appendix. ■

The only difference between the social efficiency conditions of Theorems 1 and 2 is that the social welfare test extends over the entire space $[0, \infty)$, while for poverty the test is limited to the range of potential poverty lines $[0, z^+]$. For $\gamma = 1$, Theorem 1 says that the tax reform will reduce poverty in an ethically robust manner if the $\overline{CD}$ curve of good l is higher than the $\overline{CD}$ curve of good j for every poverty line under z^+. When the range of poverty lines is unbounded, Theorem 2 extends such social efficiency to (global) welfare efficiency. For $\gamma \neq 1$, one simply translates the $\overline{CD}$ curve of good j by the economic efficiency parameter γ before checking again the ordering of the $\overline{CD}$ curves up to the maximum poverty line. Ethical robustness of poverty reduction implies that poverty will be reduced by the tax reform for all choices of poverty indices within $\Pi^s(z)$ and for all choices of poverty lines within $[0, z^+]$. Ethical robustness of welfare increase means that social welfare will be increased by the tax reform for all choices of social welfare functions in Ω^s. A tax reform is Pen efficient, Dalton efficient and Kolm efficient if (22) holds for $s = 1, 2$ or 3 respectively.

5 Critical poverty lines and efficiency ratios

The ratio $\overline{CD}_l^s(z)/\overline{CD}_j^s(z)$ of normalized consumption dominance curves can be interpreted as the distributive benefit of taxing j instead of l. To see why, assume again an increase in the tax on good j and thus a fall in the price of good l. The socially-weighted gain from this fall in q_l, as a proportion of the average welfare gain, is given by $\overline{CD}_l^s(z)$, which is therefore an indicator of the distributive benefit of taxing j. The same naturally holds for the distributive benefit of taxing l, which is given by $\overline{CD}_j^s(z)$. Denote this distributive benefit ratio as $\delta^s(z) = \overline{CD}_l^s(z)/\overline{CD}_j^s(z)$. If $\overline{CD}_j^s(z) = 0$, the relative distributive benefit of taxing j is then effectively infinite, but for tractability we will then define it as γ^{++}, which we may set to as large a finite value as we wish. $\delta^s(z)$ is thus given by:

$$\delta^s(z) = \begin{cases} \frac{\overline{CD}_l^s(z)}{\overline{CD}_j^s(z)} \text{ if } \overline{CD}_j^s(z) \neq 0 \\ \gamma^{++} \text{ if } \overline{CD}_j^s(z) = 0. \end{cases} \tag{23}$$

Note that $\delta^1(z)$ is a normalized ratio of Engel curves at income z. More precisely, it is the ratio, at z, of the share of good l over the share of good j, divided by the ratio of the average shares over the entire population.

Thus, $\delta^s(z)$ is the distributive benefit of taxing j relative to that of taxing l. Recall that γ is the economic cost of taxing j relative to that of taxing l. Comparing distributive benefit $\delta^s(z)$ to economic cost γ is crucial in determining whether a tax reform that increases t_j and decreases t_l is socially efficient. We can indeed re-write the conditions of Theorems 1 and 2 by checking whether $\delta^s(z) \geq \gamma$ for all $z \in [0, z^+]$ and for all $z \in [0, \infty)$ (equations (21) and (22)), respectively. In words, a tax reform is s-order socially efficient if its distributive benefit exceeds its economic cost over a range of alternative poverty lines. Note that when preferences are identical and homothetic, then by (19) $G_j(y) = G_l(y)$ for all y, and therefore $\delta^s(z) = 1$ whatever the values of z and s. There is then no distributive benefit to the tax reform whatever the social welfare or poverty objectives. The optimal tax system is then one in which the marginal deadweight loss of taxation is the same across the two goods.

For a given value of $\gamma = \gamma^+$, we assume from now onwards that – for some z^+ and without loss of generality – we have that

$$\overline{CD}_l^s(z) - \gamma \overline{CD}_j^s(z) \geq 0, \text{ for all } z \in [0, z^+], \tag{24}$$

with strict inequality over at least part of the range. Such curves $\overline{CD}_l^s(z)$ and $\gamma\overline{CD}_j^s(z)$ are shown on Figure 1. Condition (24) is clearly also valid for all other values of $\gamma < \gamma^+$. Hence, either (24) also holds for all superior values of $\gamma > \gamma^+$, in which case we must have that $\delta^s(z) = \gamma^{++}$ for all $z \in [0, z^+]$. Or, there exists a critical value of γ beyond which (24) does not hold everywhere over $[0, z^+]$, and this is defined by $\inf\{\delta^s(z)\,|z \in [0, z^+]\}$. Denoting this critical economic efficiency threshold as $\gamma_s(z^+)$, we have:

$$\gamma_s(z^+) = \inf\left\{\delta^s(z)\,\middle|z \in [0, z^+]\right\}. \tag{25}$$

Such a value of $\gamma_s(z^+)$ is shown in Figure 1 as that precise value of γ which makes $\overline{CD}_l^s(z)$ and $\gamma\overline{CD}_j^s(z)$ cross at $z = z^+$.

A similar exercise leads to a definition of a critical upper poverty threshold $z_s(\gamma^+)$. To see this, assume again that, for a given value of $\gamma = \gamma^+$, (24) holds for all values of z within some bottom range of $z \in [0, z^+]$, with strict inequality over at least part of the range. Then, either (24) holds for any larger value of z^+, in which case we can set $z_s(\gamma^+)$ to as large a value as we wish and denote it by z^{++}. Or, there exists a critical value of z^+ beyond which (24) does not hold anymore, and this is given by $\sup\{z\,|\delta^s(y) \geq \gamma^+, y \in [0, z]\}$. We may then define $z_s(\gamma^+)$ as:

$$z_s(\gamma^+) = \sup\left\{z\,\middle|\delta^s(y) \geq \gamma^+, y \in [0, z], z \leq z^{++}\right\}. \tag{26}$$

Figure 1 shows such a $z_s(\gamma)$ as the first crossing point of $\overline{CD}_l^s(z)$ and $\gamma\overline{CD}_j^s(z)$. Equations (25) and (26) say that increasing t_j and decreasing t_l is socially efficient so long as γ and z are not allowed to exceed certain critical thresholds, which depend on the order of ethical dominance s. For a given γ^+ and z^+, $z_s(\gamma^+)$ and $\gamma_s(z^+)$ give respectively the critical upper poverty line and the critical economic efficiency ratio up to which the tax reform is necessarily s-order poverty efficient.

$z_s(\gamma^+)$ also gives the first poverty line (starting from $z = 0$) at which a poverty analyst using an $FGT^{s-1}(z)$ poverty index would be exactly indifferent to the tax reform when the MCPF ratio is given by γ^+. At all lower poverty lines, the FGT poverty analyst would choose increasing t_j, and at poverty lines just higher than $z_s(\gamma^+)$ he would prefer decreasing t_j. Alternatively, $\gamma_s(z^+)$ gives the MCPF ratio for which a $FGT^{s-1}(z^+)$ analyst would be exactly indifferent between reallocating or not the burden of indirect taxation across the goods j and l. At lower values of γ, the FGT analyst prefers taxing j further; at higher values, he prefers taxing l further.

6 Discussion

A tax reform cannot be Pen efficient when $\gamma \geq 1$, as Lemma 9 shows in the appendix. Pen-efficient tax reforms are, however, theoretically possible so long as $\gamma < 1$. Yitzhaki and Thirsk (1990) and Yitzhaki and Slemrod (1991) find that if γ is larger than one, a tax reform cannot be Dalton efficient, a result which is also shown in Lemma 10 in the Appendix. The proof of Proposition 2 also shows that $\gamma_s(\infty) \leq 1$ for any $s = 3, 4, ...$ Thus, social efficiency of any ethical order requires economic efficiency.

A tax reform may, however, be poverty efficient at any order – even at the first – even when $\gamma > 1$. For first- and second-order poverty efficiency, the excess burden or economic efficiency cost of the reform then has to be paid by those households whose income is above the maximum poverty line z^+, and to whose change in welfare poverty analysts are ethically indifferent. More surprisingly perhaps, for ethical orders $s = 3, 4, ...$, a tax reform may be poverty efficient when $\gamma > 1$ even when $z^+ > a$ (which says that all can then be considered poor, though *not* equally poor). We sum up these results in the following remark:

Remark 3 *Social efficiency of any order requires economic efficiency. Economic efficiency is not, however, needed for poverty efficiency; for* $s > 2$ *, this is true even in cases in which everyone may be considered poor.*

If a tax reform is s-order poverty efficient up to some z^+, then it is also poverty efficient up to z^+ at the $s+1$ order (Lemma 11). Lemma 11 also says that if $\gamma_s(z^+) < \gamma^{++}$ and if there is strict dominance over a bottom range of z, then $\gamma_{s+1}(z^+) > \gamma_s(z^+)$, meaning that, as the order of ethical dominance increases, economic efficiency becomes less constraining. Increasing the order of dominance also increases the range of poverty lines over which a reform is poverty efficient. Indeed, if $z_s(\gamma^+) < z^{++}$ and if there is strict dominance over a bottom range of z, then $z_s(\gamma^+) < z_{s+1}(\gamma^+)$ (see Lemma 12). Finally, if there is poverty efficiency over a bottom range that extends to z^+, with strict dominance over at least part of that range, then for any other finite threshold z^+, there will also be poverty efficiency for a sufficiently large order s (this is by Lemma 1 of Davidson and Duclos (2000)).

These relationships are shown on Figure 2 using the distributive benefit ratios of equation (23). When a reform's distributive benefit $\delta^s(z)$ over $[0, z^+]$ exceeds its economic cost γ, the reform is deemed poverty efficient. At order

$s = 1$, this is the case in Figure 2 for all γ up to $\gamma_1(z^+)$. It is also necessarily true for all higher orders $s = 2, 3, ...$ It can also be seen that the distributive benefit increases with s over $[0, z^+]$, which then graphically means that $\gamma_s(z^+)$ also rises with s (*e.g.*, $\gamma_3(z^+) > \gamma_1(z^+)$ in Figure 2). Alternatively, for a given γ^+, the critical poverty line $z_s(\gamma^+)$ (up to which the distributive benefit is higher than the economic cost) increases with s; on Figure 2, this is shown by $z_3(\gamma^+) > z_2(\gamma^+) > z_1(\gamma^+)$.

The Dalton efficiency condition of Theorem 2 is equivalent for $s = 2$ to a condition derived in Yitzhaki and Slemrod (1991) and Yitzhaki and Thirsk (1990). Their approach, however, is different. Their condition must be tested over the range of percentiles $[0, 1]$, while ours is defined over an income range. To see this difference more clearly, define the concentration curve for good k as $\overline{C}_k^2(p)$:

$$\overline{C}_k^2(p) = \frac{1}{X_k(q)} \int_0^{F^{-1}(p)} \overline{x}_k(y, q) dF(y) \tag{27}$$

where $F^{-1}(p)$ is the (left) inverse of the marginal income distribution function of incomes, $F^{-1}(p) = \inf\{s > 0 \,|F(s) \geq p\}$. $F^{-1}(p)$ is often called the p-quantile, that is, roughly speaking, the income of the individual whose rank is p. Hence, $\overline{CD}_k^2(z) = \overline{C}_k(F(z))$. There thus exists a parallel between the $\overline{CD}_k^2(z)$ and the usual concentration curves. The $\overline{CD}_k^2(y)$ curves gives the share in good k of those below y, whilst $\overline{C}_k^2(p)$ yields the share in k of those with rank p or below. Concentration curves thus act as *dual* $\overline{CD}$ curves, just as it is well known that Generalized Lorenz curves (see Shorrocks (1983)) act as *dual* $D^2(z)$ dominance curves. With the dual $\overline{C}_k^2(p)$ curves, Dalton efficiency is tested by checking $\overline{C}_l^2(p) - \gamma\overline{C}_j^2(p) \geq 0$, $\forall p \in [0, 1]$.

We may also define a dual first-order CD curve, denoted as $\overline{C}_k^1(p)$:

$$\overline{C}_k^1(p) = \frac{\overline{x}_k(F^{-1}(p), q)}{X_k(q)}, \quad p \in [0, 1]. \tag{28}$$

These curves simply show the expected consumption of good k (relative to average consumption) along Pen's income parade. Pen efficiency can be tested by verifying whether $\overline{C}_l^1(p) - \gamma\overline{C}_j^1(p) \geq 0$, $\forall p \in [0, 1]$. As can be easily seen, this is equivalent to checking $\overline{CD}_l^1(y) - \gamma\overline{CD}_j^1(y) \geq 0$, $\forall y \in [0, a]$.

Similarly, first-order poverty efficiency can be tested by checking whether $\overline{C}_l^1(p) - \gamma \overline{C}_j^1(p) \geq 0$, $\forall p \in [0, F(z^+)]$. Using primal as opposed to dual curves has, however, the advantage of simplifying testing procedures for dominance orders higher than 2.[9]

Two additional points are worth discussing. Although notationally and expositionally more complicated, it would not be analytically much more difficult to allow for general equilibrium effects from changes in t. This would involve considering changes in producer prices, with effects on the welfare of producers, which would typically feed into changes in the consumer and the producer prices of goods other than those whose tax rates are changed by the government. The analysis would then take into account the sum of the welfare effects of the marginal changes in the various consumer and producer prices. We hope to consider such a generalization in future work.

The tools developed above are in principle strictly limited to the analysis of infinitesimal price changes. For changes in tax rates that are not infinitely small, the impact of a change in t_k on an exact measure of individual welfare is not precisely given by the pre-reform observed demand of good k times the change in t_k. The proportional difference between the exact measure and the inexact measure that we use is approximately equal to one half the compensated price elasticity of good k times dt_k/dq_k. Hence, for a compensated price elasticity of 1, a 5% increase in the price of a good k leads approximately to a 2.5% error in the estimate of consumer welfare change when one uses the pre-reform consumption of k to compute that estimate. For many purposes, this error would seem relatively small.

7 Pareto efficient tax reforms

We mentioned above that Pen efficiency implies robustness of social welfare increase over all social welfare functions that are increasing and anonymous (or symmetric) in real incomes. Pareto efficiency does not, however, impose anonymity as a property of the social evaluation exercise. This difference between these two efficiency concepts may be subtle, but it is fundamental to the usefulness of practical searches for socially efficient tax reforms.

[9] Recall that the concentration curves used by Yitzhaki and Slemrod (1991) and Yitzhaki and Thirsk (1990) were used solely for second-order welfare efficiency. Note also that – following Zoli (1999) and Aaberge (2001) – one could probably link some yet-to-defined higher-order concentration curves to some of their "rank-order" ethical principles.

A tax reform is Pareto efficient if it decreases no one's real income. Under a marginal tax reform for which $dq_j = -\gamma\left(X_l/X_j\right)dq_l$, the impact on one's real income (with nominal income y and preferences θ) is given by (consider (3)):

$$-x_j(y,\theta)dq_j + \frac{X_j}{\gamma X_l}x_l(y,\theta)dq_j. \tag{29}$$

This is non-negative for $dq_j > 0$ if and only if $\overline{x}_l(y,\theta) - \gamma\overline{x}_j(y,\theta) \geq 0$. Thus:

Definition 4 *A tax reform is Pareto efficient if and only if*

$$\begin{aligned} \overline{x}_l(y,\theta,q) - \gamma\overline{x}_j(y,\theta,q) &\geq 0, \text{ for all } y \in [0,\infty) \text{ and for all } \theta \in \Omega \\ \text{for which } dF(y) &> 0 \text{ and } dF(\theta|y) > 0. \end{aligned} \tag{30}$$

Define now the "distributive benefit" of the tax reform for someone at z and with preferences θ as

$$\delta^0(z,\theta) = \begin{cases} \frac{\overline{x}_l(z,\theta)}{\overline{x}_j(z,\theta)} \text{ if } \overline{x}_j(z,\theta) \neq 0 \\ \gamma^{++} \text{ if } \overline{x}_j(z,\theta) = 0. \end{cases} \tag{31}$$

Denote by $\gamma_0(z^+)$ the maximum ratio of the MCPF for which the tax reform is Pareto efficient. This is formally defined by

$$\gamma_0(z^+) = \inf\left\{\delta^0(y,\theta)\left|y \in [0,z^+], \theta \in \Omega, dF(y) > 0,\ dF(\theta|y) > 0\right.\right\}. \tag{32}$$

Let also $z_0(\gamma^+)$ be the critical poverty line up to which a tax reform is Pareto efficient:

$$z_0(\gamma^+) = \sup\left\{z\left|\delta^0(y,\theta) \geq \gamma^+, y \in [0,z], \theta \in \Omega, dF(y,\theta) > 0, z \leq z^{++}\right.\right\}. \tag{33}$$

Lemma 13 in the appendix shows that $\gamma_0(z^+) \leq \gamma_1(z^+)$, with strict inequality if there is heterogeneity in the ratio of Engel curves for goods l and j at each

$y \in [0, z^+]$. Lemma 14 further shows that $z_0(\gamma^+) \leq z_1(\gamma^+)$, with strict inequality if $z_1(\gamma^+) < z^{++}$ and if there is heterogeneity in the ratio of Engel curves at $z_1(\gamma^+)$. These are just formal ways of saying that Pareto-efficient reforms are theoretically more difficult to identify than first-order socially efficient ones. It also implies that the implementation of a first-order efficient tax reform with a MCPF ratio equal to $\gamma_1(z^+)$ will usually generate losers among those whose income is below z^+.

An arguably more important issue is whether in practice $\gamma_0(z)$ is significantly lower than $\gamma_1(z^+)$, or whether $z_0(\gamma^+)$ is substantially lower than $z_1(\gamma^+)$. We can expect condition (30) to be very restrictive empirically (as has been argued before[10]). In fact, it would not be misleading to suggest that a search for Pareto-efficient tax reforms will normally be doomed to failure. This is due to the considerable heterogeneity of preferences typically found in the observed consumption patterns of a real population of individuals. With Pareto efficiency as a social welfare constraint, it will be very difficult to identify any socially-efficient movement away from the current tax system, thus giving a sort of "vetoing status" to existing tax systems, whatever the defects of these systems may be. Furthermore, the identification of Pareto-efficient tax reforms requires the observation of consumption patterns that are free of measurement errors[11]. This is less of a problem for Pen efficiency since it is the expected value of consumption patterns that matters, and in computing those expected values the measurement errors are (at least partially) averaged out. Whether the concept of Pen efficiency eases the search for socially-efficient reforms is ultimately, of course, an empirical issue, to which we will revert later.

8 Estimation and inference

We now turn to the estimation of some of the analytical tools developed above. For this, we suppose for expositional simplicity that we dispose of a sample of N independently and identically distributed observations[12], and that the pre-reform income and consumption of goods j and l for observation

[10] See for instance Ahmad and Stern (1984 and 1991) and the comments *inter alia* in Yitzhaki and Thirsk (1990) and Yitzhaki and Slemrod (1991).

[11] See a discussion of this in Ahmad and Stern (1984), p.290.

[12] The analytical results can be extended to account for complex multi-stage sampling designs.

i $(i = 1, ..., N)$ are denoted by y_i, x^i_j and x^i_l, respectively. For $s \geq 2$, the CD curves can then be estimated by the natural estimator

$$\widehat{CD}^s_k(z) = \frac{1}{(s-2)!}\frac{1}{N}\sum_{i=1}^{N} x^i_k \left(z - y_i\right)^{s-2}_+ , \tag{34}$$

where $f_+ = \max(0, f)$. The estimator $\widehat{CD}^s_l(z) - \gamma\widehat{CD}^s_j(z)$ is given analogously. Note that $(s-1)!z^{1-s}\widehat{CD}^s_k(z)dq_k$ is an estimator of the impact of a marginal change dq_k on the $FGT^{s-1}(z)$ index. The asymptotic sampling distribution of $\widehat{CD}^s_k(z)$ for $s \geq 2$ is given in Theorem 5.

Theorem 5 *Let the second population moment of* $x_k(y, \theta)\left(z - y\right)^{s-2}_+$ *be finite. Then, for* $s \geq 2$, $N^{0.5}\left(\widehat{CD}^s_k(z) - CD^s_k(z)\right)$ *is asymptotically normal with mean zero and with asymptotic variance given by:*

$$\begin{aligned} & \lim_{N\to\infty} N \cdot var\left(\widehat{CD}^s_k(z) - CD^s_k(z)\right) \\ = \;& (s-2)!^{-2} E\left[\left(x_k(y,\theta)(z-y)^{s-2}_+\right)^2\right] - CD^s_k(z)^2 . \end{aligned} \tag{35}$$

The asymptotic distribution of $\widehat{CD}^s_l(z) - \gamma\widehat{CD}^s_j(z)$ follows directly from Theorem 5 simply by replacing $\widehat{CD}^s_k(z)$ by $\widehat{CD}^s_l(z) - \gamma\widehat{CD}^s_j(z)$, $CD^s_k(z)$ by $CD^s_l(z) - \gamma CD^s_j(z)$, and $x_k(y,\theta)(z-y)^{s-2}_+$ by $\left(x_l(y,\theta) - \gamma x_j(y,\theta)\right)(z-y)^{s-2}_+$. Normalized $\overline{CD}$-curves can be defined as for $\widehat{\overline{CD}}^s_k(z)$ in (34) by dividing x^i_k by the estimate of average consumption $\widehat{X}_k = \frac{1}{N}\sum_{i=1}^{N} x^i_k$. The sampling distribution of $\widehat{\overline{CD}}^s_k(z)$ can then be obtained by linearizing $\widehat{\overline{CD}}^s_k(z)$ with respect to $\widehat{CD}^s_k(z)$ and $\widehat{\overline{x}}_k$.[13] The asymptotic distribution of $\widehat{\overline{CD}}^s_k(z)$ is subsequently given by Theorem 5 by replacing $\widehat{CD}^s_k(z)$ by $\widehat{\overline{CD}}^s_k(z)$, $CD^s_k(z)$ by $\overline{CD}^s_k(z)$, and the limiting variance by

$$\lim_{N\to\infty} N \cdot var\left(\widehat{\overline{CD}}^s_k(z) - \overline{CD}^s_k(z)\right) \tag{36}$$

$$= E\left[\overline{x}_k(y,\theta)^2\left((s-2)!^{-1}(z-y)^{s-2}_+ - \overline{CD}^s_k(z)\right)^2\right]. \tag{37}$$

[13] See, for instance, Rao (1973). This first-order approximation is often called the "delta method".

An exactly similar procedure can be applied for $\widehat{CD}_l^s(z) - \gamma \widehat{CD}_j^s(z)$. Multiplying $\widehat{CD}_l^s(z) - \gamma X_l \widehat{CD}_j^s(z)/X_j$ by $(s-1)! z^{1-s} dq_l$ yields an estimator of the marginal impact of the reform on the $FGT^{s-1}(z)$ index.

For $s = 1$, recall that we have that

$$CD_k^1(z) = x_k(z)\, f(z). \tag{38}$$

Testing first-order social efficiency therefore requires an estimator of the product of the expected consumption of good k times the density of income at the poverty line, $f(z)$. For this, we can use non-parametric estimation procedures, using for instance a kernel estimator defined such as

$$\widehat{CD}_k^1(z) = \frac{1}{N} \sum_{i=1}^{N} \kappa_h(z - y_i)\, x_k^i, \tag{39}$$

where $\kappa_h(u) = h^{-1}\kappa(u/h)$, $\int \kappa(u) du = 1$, $\int u\kappa(u)\, du = 0$ (for symmetry), and $\int u^2 \kappa(u)\, du = c_\kappa$. In the illustration below, we choose a Gaussian form for $\kappa(u)$,

$$\kappa(u) = \frac{e^{-0.5u^2}}{\sqrt{2\pi}}, \tag{40}$$

but other kernel functional forms could also be used. In the illustration, we choose h using the cross-validation method, which is asymptotically optimal (see Härdle (1990), Theorem 5.1.1). Theorem 6 then gives the asymptotic sampling distribution of $\widehat{CD}_k^1(z)$.

Theorem 6 *Let i) $\int \kappa(u)^2 du$ exists, ii) $h \sim N^{-0.2}$, iii) $CD_k^1(y)$ be twice differentiable in y at $y = z$, iv) $f(z) > 0$, and v) $c_k(z) = \int_\Omega x_k(z,\theta)^2 dF(\theta|z)$ be continuous at z. Then, $(Nh)^{0.5}\left(\widehat{CD}_k^1(z) - CD_k^1(z) - h^2 B_k(z)\right)$ is asymptotically normal with mean 0 and limiting variance $V_k(z)$, where $B_k(z) = 0.5\, c_\kappa CD_k^{1''}(z)$ and $V_k(z) = f(z) c_k(z) \int \kappa(u)^2 du$.*

The asymptotic distribution of $\widehat{CD}_l^1(z) - \gamma \widehat{CD}_j^1(z)$ follows again directly from Theorem 6 by appropriate substitutions. $(Nh)^{0.5} X_k V_k(z)^{-0.5}$

$\left(\widehat{\overline{CD}}_k^1(z) - \overline{CD}_k^1(z) - h^2 B_k(z)/X_k\right)$ also has a limiting standard normal distribution[14].

For $s = 1$, we can therefore test social efficiency using:

$$\widehat{d}_1(z,\gamma) = \widehat{\overline{CD}}_l^1(z) - \gamma\widehat{\overline{CD}}_j^1(z) - h^2(B_l(z)/X_l - \gamma B_j(z)/X_j) \tag{41}$$

and, for $s \geq 2$, we can use:

$$\widehat{d}_s(z,\gamma) = \widehat{\overline{CD}}_l^s(z) - \gamma\widehat{\overline{CD}}_j^s(z) \tag{42}$$

with $d_s(z,\gamma) = \overline{CD}_l^s(z) - \gamma\overline{CD}_j^s(z)$ the corresponding population values. The terms needed to carry out statistical inference are either constants (c_κ and $\int \kappa(u)^2 du$) or can be readily estimated consistently in a distribution-free manner (this is the case, for instance, of $E\left[\left(x_k(z-y)_+^{s-2}\right)^2\right]$, $CD_k^s(z)^2$, $CD_k^{1''}(z)$, $f(z)$ and $c_k(z)$). Note, however, that it is usual to consider (and to find) the bias terms $B_k(z)$ and $B_k(z)/X_k$ to be of negligible practical importance[15], and we also make this assumption in the illustration below and in deriving the limiting variance of the distribution of the estimators of $z_1(\gamma^+)$ and $\gamma_1(z^+)$. These are given (for $s = 1, 2, ...$) respectively by $\widehat{z}_s$ and $\widehat{\gamma}_s$ and are defined as:

$$\widehat{\gamma}_s(z^+) = \sup\left\{\gamma \left| \widehat{d}_s(y,\gamma) \geq 0, y \in [0, z^+], \gamma \leq \gamma^{++}\right.\right\}$$

and

$$\widehat{z}_s(\gamma^+) = \sup\left\{z \left| \widehat{d}_s(y,\gamma^+) \geq 0, y \in [0, z], z \leq z^{++}\right.\right\}. \tag{43}$$

Note that $\widehat{z}_s(\gamma^+)$ are estimators of the point at which $\overline{CD}$ curves first cross. $\widehat{z}_1(\gamma^+)$ is, in particular, an estimator of the first point at which kernel regression curves intersect. Denote

$$d_s^z(z,\gamma) \equiv \left.\frac{\partial d_s(y,\gamma)}{\partial y}\right|_{y=z}.$$

[14] This is in part because $\widehat{X}_k - X_k = O(N^{-0.5})$ is of smaller order than $\widehat{CD}_k^1(z) - E\left[\widehat{CD}_k^1(z)\right] = O((Nh)^{-0.5})$.

[15] This is particularly true in the study of consumption data, where the second order derivative of expected consumption at z, $CD_k^{1''}(z)$, may be expected to be small. For more on this, see for instance Härdle (1990), p.101.

It can easily be checked that $d_s^z(z,\gamma) = d_{s-1}(z,\gamma)$ for $s > 1$, which is thus easily estimated. For $s = 1$, $d_1^z(z,\gamma)$ is given by:

$$d_1^z(z,\gamma) = \left. \frac{\partial \underset{\theta}{E}\left[\overline{x}_j\left(y,\theta,q^R\right) - \gamma\overline{x}_l\left(y,\theta,q^R\right)\right]}{\partial y} \right|_{y=z}, \tag{44}$$

which can again be estimated consistently, using in (39) derivatives of the kernel functions $\kappa(z,y_i,\theta)$ instead of the functions themselves. Theorems 7 and 8 then give the asymptotic sampling distribution of $\widehat{z}_s(\gamma^+)$ and $\widehat{\gamma}_s(z^+)$ respectively.

Theorem 7 *Assume that there exists $z_s(\gamma^+) < z^{++}$ such that $d_s(z_s(\gamma^+),\gamma^+) = 0$, $d_s(z,\gamma^+) > 0$ for all $z < z_s(\gamma^+)$, and $d_s^z(z_s(\gamma^+),\gamma^+) < 0$. For $s = 2, 3, ...,$ let the second population moment of $(z_s(\gamma^+) - y)_+^{s-2}(x_j(y,\theta) - \gamma^+ x_l(y,\theta))$ be finite. Then, for $s \geq 2$, $N^{0.5}(\widehat{z}_s(\gamma^+) - z_s(\gamma^+))$ is asymptotically normal with mean zero and with asymptotic variance given by*

$$\begin{aligned}
& \lim_{N\to\infty} N \cdot var\left(\widehat{z}_s(\gamma^+) - z_s(\gamma^+)\right) \\
= \;& d_s^z(z_s(\gamma^+),\gamma^+)^{-2} E\left[\left\{(s-2)!^{-1}\overline{x}_l(y,\theta)\left((z_s(\gamma^+) - y)_+^{s-2} - \overline{CD}_l^s\left(z_s(\gamma^+)\right)\right)\right.\right. \\
& \left.\left. -\gamma^+\overline{x}_j(y,\theta)\left((s-2)!^{-1}(z_s(\gamma^+) - y)_+^{s-2} - \overline{CD}_j^s\left(z_s(\gamma^+)\right)\right)\right\}^2\right].
\end{aligned} \tag{45}$$

For $s = 1$, assume that conditions analogous to those of Theorem 6 for $\widehat{CD}_k^1(z_1(\gamma^+))$ hold for $\widehat{CD}_l^1(z_1(\gamma^+)) - \gamma^+\widehat{CD}j^1(z_1(\gamma^+))$. $(Nh)^{0.5}(\widehat{z}_1(\gamma^+) - z_1(\gamma^+))$ is then asymptotically normal with mean zero and with asymptotic variance given by

$$\begin{aligned}
& \lim_{N\to\infty} Nh \cdot var\left(\widehat{z}_1(\gamma^+) - z_1(\gamma^+)\right) \\
= \;& d_1^z(z_1(\gamma^+),\gamma^+)^{-2} f(z_1(\gamma^+)) c_{lj}(z_1(\gamma^+),\gamma^+) \int \kappa(u)^2\, du
\end{aligned} \tag{46}$$

where $c_{lj}(z,\gamma) = \int_\Omega \left(\overline{x}_l(z,\theta) - \gamma\overline{x}_j(z,\theta)\right)^2 dF(\theta|z)$.

Theorem 8 *Assume that $CD_l^s(y) > 0$ over some interval $y \in [0, z^+]$. Let $d_s(y,\gamma) \geq 0$ for all $\gamma < \gamma_s(z^+)$ and for all $y \in [0, z^+]$, and let $\varsigma = \sup\left\{z\,|d_s(z,\gamma_s(z^+)) = 0, z \in [0, z^+]\right\}$. For $s = 2, 3...,$ let the second population moment of $(\varsigma - y)_+^{s-2}(x_j(y,\theta) - \gamma x_l(y,\theta))$ be finite. Then, $N^{0.5}(\widehat{\gamma}_s(z^+) - \gamma_s(z^+))$*

is asymptotically normal with mean zero and with asymptotic variance given by

$$\begin{aligned}
&\lim_{N\to\infty} N \cdot var\left(\widehat{\gamma}_s(z^+) - \gamma_s(z^+)\right) \\
= \; & \overline{CD}_j^s(\varsigma)^{-2} E\left[\left\{\overline{x}_l(y,\theta)\left((s-2)!^{-1}(\varsigma - y)_+^{s-2} - \overline{CD}_l^s(\varsigma)\right)\right.\right. \qquad (47)\\
& \left.\left. -\gamma^+ \overline{x}_j(y,\theta)\left((s-2)!^{-1}(\varsigma-y)_+^{s-2} - \overline{CD}_j^s(\varsigma)\right)\right\}^2\right]. \qquad (48)
\end{aligned}$$

For $s = 1, (Nh)^{0.5}\left(\widehat{\gamma}_s(z^+) - \gamma_s(z^+)\right)$ *is asymptotically normal with mean zero and with asymptotic variance given by*

$$\begin{aligned}
&\lim_{N\to\infty} Nh \cdot var\left(\widehat{\gamma}_s(z^+) - \gamma_s(z^+)\right) \\
= \; & \overline{CD}_j^s(\varsigma)^{-2} f(\varsigma) c_{lj}(\varsigma, \gamma^+) \int \kappa(u)^2 \, du. \qquad (49)
\end{aligned}$$

with $c_{lj}(z,\gamma) = \int_\Omega \left(\overline{x}_l(z,\theta) - \gamma \overline{x}_j(z,\theta)\right)^2 dF(\theta|z)$.

9 Empirical illustration

This section briefly applies the above normative and statistical tools to household-level data from Mexico's 1996 ENIGH, a nationally representative survey with detailed income and consumption modules. Vicente Fox, the newly elected President, has proposed in April 2001 a wide-ranging tax reform whose objectives are to reduce inefficiencies in the tax system while also raising additional revenues in order to reduce the budget deficit. There are three main elements in the proposed reform: a) the personal direct income tax brackets would be simplified and the marginal tax rates would be reduced; b) the exemptions and zero ratings for the Value Added Tax (VAT) for some categories of expenditures would be eliminated; and c) the resources devoted to targeted social programs benefiting the poor would be increased to offset the negative impact on the poor of the reform of the VAT. It is beyond the scope of this paper to assess whether this proposed global tax reform will be socially efficient, but the tools presented above can be readily illustrated if we focus on its VAT component.

The main potential negative impact for the poor of the proposed reform comes from the termination of the VAT exemption for food expenditures. The rationale for the exemption is apparent in Figure 3, which provides (normalized) expected food and non-food expenditures $\overline{CD}_{food}^1(z)$ and

$\overline{CD}^1_{non-food}(z)$ at different incomes z. Per capita incomes (on the horizontal axis) have been normalized by cost of living indices (the regional poverty lines used in the latest poverty assessment for Mexico completed at the World Bank), so that cost-of-living differences between urban and rural areas are taken into account. A value of one indicates that a household is at the level of the urban/rural poverty line. With these poverty lines, 60.9 percent of the population is poor (those with per capita income below $z = 1$). The two $\overline{CD}^1(z)$ curves cross at $\widehat{z}_1(1) = 1.793$, which is also the critical first-order poverty line $z_1(\gamma)$ shown in the bottom panel of Table 1 for $\gamma = 1$. Note that the standard error of the sampling distribution of $\widehat{z}_1(\gamma)$ is estimated to be 0.099, which implies that a 95% confidence interval for the true value $z_1(\gamma)$ would be $[1.60, 1.98]$. Simply stated, this means that if $\gamma = 1$, we can be 95% certain that for any poverty line below 1.60 it is first-order poverty efficient to implement a balanced-budget indirect tax reform by reducing at the margin taxes on (or providing a subsidy for) food expenditures and increasing taxes on non-food expenditures.

However, and as expected, this reform is not Pen efficient for $\gamma^+ = 1$ since we do find a critical poverty line at which the two curves intersect. Furthermore, while a first-order poverty efficient (including headcount-reducing) food/non-food tax reform is feasible for poverty lines below $z_1(1)$, a substantial share of the poor would lose from such a reform. Given the average food and non-food expenditure shares, with $\gamma = 1$ a one percentage point tax reduction for food must be compensated by a 0.340 percentage point tax increase on non-food expenditures for budget neutrality. Hence, any poor household whose food expenditure share is below one fourth will lose from such a reform. The relative importance of these losers appears in Figure 4, which gives the cumulative population share and the cumulative share of losers as functions of the poverty line. The ratio of the two cumulative shares represents the share of the population below a given poverty line that loses from the reform. It turns out that 19.0 percent of the population below $z = 1$ and 14.3 percent of the population below $z = 0.5$ would loose from this hypothetical tax reform. Hence, this reform is clearly not Pareto efficient, and would not be so even if we were to censor the assessment of its impact at a poverty line. In fact – and as anticipated in Section 7 – Pareto-efficient food and non-food tax reforms are virtually impossible, whatever the MCPF economic efficiency parameter. This is due to the fact that for both food and non-food expenditures, there are households with either zero or very low expenditure levels across the income distribution, so that $\widehat{\gamma}_0(z^+)$ is zero for virtually any z^+.

Taxing non-food expenditures and reducing taxes (or providing subsidies) for food expenditures is Dalton efficient for a wide range of values of γ. This is demonstrated in Figure 5 – $\overline{CD}^2_{food}(z)$ is everywhere above $\overline{CD}^2_{non-food}(z)$. The top panel of Table 1 gives the values of the $\overline{CD}^2_s$ curves for $z = 0.5$, $z = 1$, and $z = 2$, as well as the difference between the curves (all differences are statistically greater than zero). For $s = 2$, these curves represent the cumulative shares of food and non-food expenditures accounted for by those with per capita income below a certain level. For example, the population below $z = 1$ accounts for 41.0 percent of total food expenditures and 24.7 percent of total non-food expenditures. Figure 5 also provides the distributive benefit ratio $\delta^2(z)$. As long as the economic cost γ of taxing food items relative to non-food items is below $\delta^2(z)$, taxing non-food items to give relief to food items is beneficial.

The critical efficiency ratios $\gamma_s(z^+)$ under which the tax reform is Dalton efficient are provided in the middle panel of Table 1. For $s = 2$, $\gamma_2(z^+)$ is equal to, respectively, 2.021 for $z^+ = 0.5$, 1.657 for $z^+ = 1$, and 1.390 for $z^+ = 2$. If, following standard practice in Latin America, we consider those with *per capita* income below half the poverty line as extreme poor, the tax reform would reduce all distributive-sensitive poverty indices for the extreme poor even if the economic cost of non-food taxation were 100 percent higher than that of food taxation.

The bottom panel of Table 1 also gives the critical poverty lines $\widehat{z}_s(\gamma^+)$ under which the tax reform would remain efficient for poverty reduction with various MCPF ratios. For example, with $\gamma^+ = 1.5$, the tax reform is second-order poverty reducing up to a poverty line of 0.752 (with a standard error of 0.03). For $\gamma^+ = 1$, there is no critical poverty line for $s = 2$ or higher. This is due to the fact that $\overline{CD}^2_{food}(z)$ and $\overline{CD}^2_{non-food}(z)$ do not intersect. Remember also that when multiplied by $(s-1)!/z^{s-1}$, $\overline{CD}^s(z)$ gives the impact on the $FGT^{s-1}(z)$ poverty indices of a marginal change in the price of a good or a marginal increase in the tax on that good. Since the differences between the $\overline{CD}^2$ curves are statistically significant for the values of z considered in Table 1, a tax reform would lead to a statistically significant reduction in the poverty gap whatever these values of z. This does not necessarily mean that universal food subsidies are the best policy option for poverty reduction. In Mexico, universal subsidies for tortillas have recently been terminated, with the savings used for better-targeted social programs, including school stipends targeted to poor children living in poor rural areas. Still, within the options provided by indirect tax reforms, it may be Dalton efficient to increase the tax on non-food expenditures while providing subsidies for food expenditures.

Figures 3 to 5 are provided for broad food and non-food expenditure aggregates, but the tools can equally well be applied to more specific items. In Figure 6, we compare the $\overline{CD}^2$ curves for a mixed bundle of food (including baby food, packaged food, deserts, drinks, and food consumed away from the home) and pasteurized milk. The estimated curves cross for a value of z slightly larger than the reference poverty line $z = 1$, but Table 2 (which is analogous to Table 1) indicates that the difference between the two curves is not statistically significant at $z = 1$ (and, similarly, that $\gamma_2(z^+ = 1)$ is not statistically larger than one). Still, since the $\overline{CD}^2$ curve for milk is below that for the mixed bundle up to $\widehat{z}_2(\gamma = 1) - 1.96 * 0.212 = 0.65$ (allowing for a 95% confidence interval), it is certainly feasible with $\gamma = 1$ to reduce extreme poverty for all distributive-sensitive poverty indices by taxing milk and providing a subsidy for the items in the mixed-food bundle. Yet, since the second-order CD curves cross, this is not a Dalton-efficient tax reform for $\gamma = 1$. It can be shown that this is also not a Kolm-efficient tax reform for $\gamma = 1$ either, since the $\overline{CD}^3$ curves intersect as well, albeit at a higher value of z (as predicted by Lemma 12).

As in Table 1, Table 2 provides the critical efficiency ratios $\gamma_2(z^+)$ and the critical poverty lines $\widehat{z}_s(\gamma^+)$ under which taxation of pasteurized milk and subsidies for the items in the mixed bundle is socially efficient. As s increases, economic efficiency is clearly less of a constraint. This is apparent in the fact that $\gamma_s(z^+)$ increases with s (as anticipated by Lemma 11). Again, building up the ethical content of the classes of poverty indices considered (*i.e.*, increasing s) increases the range of poverty lines and/or economic efficiency ratios over which the reform can confidently be deemed good for poverty reduction.

10 Conclusion

This paper shows how one can use simple Consumption Dominance curves to assess the social efficiency of indirect marginal tax reforms. The methods are similar in spirit to checking for non-intersecting (second-order) concentration curves, but they are considerably more general in that they enable the analyst to choose the order of ethical dominance in which he is interested and to censor individual welfare at some upper bound of poverty lines if so desired.

The proposed graphical tools have considerable normative appeal, in that they may be used to determine whether commodity-tax changes can be deemed to improve social welfare or decreasing poverty for large classes of social welfare and poverty indices and for broad ranges of poverty lines. They also provide detailed and useful descriptive information on the distribution of expenditures across the entire income distribution. The paper further proposes estimators of critical poverty lines and economic efficiency ratios which can be used to characterize socially-efficient tax reforms, and also derives their sampling distribution. The methodology is illustrated using Mexican data to investigate whether some aspects of a recently proposed reform of the Mexican VAT system can be considered socially efficient.

References

[1] Aaberge, R. (2001), Ranking Intersecting Lorenz Curves, Mimeo, Research Department, Statistics Norway.

[2] Ahmad, E. and N.H. Stern (1984), The Theory of Reform and Indian Indirect Taxes, *Journal of Public Economics*; 25, 259-298.

[3] Ahmad, E. and N.H. Stern (1991), *The Theory and Practice of Tax Reform in Developing Countries*, Cambridge University Press, Cambridge.

[4] Besley, T. and R. Kanbur (1988), Food Subsidies and Poverty Alleviation, *The Economic Journal*, 98, 701-719.

[5] Blackorby, C. and D. Donaldson (1978), Measures of Relative Equality and Their Meaning in Terms of Social Welfare, *Journal of Economic Theory*, 18, 59-80.

[6] Chakravarty, S.R. (1983), A New Index of Poverty, *Mathematical Social Sciences*, 6, 307-313.

[7] Christiansen, V. and E.S. Jansen (1978), Implicit Social Preferences in the Norwegian System of Indirect Taxation, *Journal of Public Economics*; 10, 217-245.

[8] Clark, S., R. Hemming and D. Ulph (1981), On Indices for the Measurement of Poverty, *Economic Journal*, 91, 515-526.

[9] Dasgupta, P., A.K. Sen and D. Starrett (1973), Notes on the Measurement of Inequality, *Journal of Economic Theory*; 6, 180-187.

[10] Davidson, R. and J.Y. Duclos (2000), Statistical Inference for Stochastic Dominance and the for the Measurement of Poverty and Inequality, *Econometrica*, 68, 1435-1465.

[11] Davies, J. and M. Hoy (1994), The Normative Significance of Using Third-Degree Stochastic Dominance in Comparing Income Distributions, *Journal of Economic Theory*, 64, 520-530.

[12] Deaton, A. S. (1977), Equity, Efficiency, and the Structure of Indirect Taxation, *Journal of Public Economics*, 8, 299-312.

[13] Donaldson, D. (1992), On the aggregation of money measures of well-being in applied welfare measures, *Journal of Agricultural and Resource Economics*, 17, 88-102.

[14] Duclos, J.-Y. and P. Makdissi (2000), Welfare, Inequality and Poverty: A Unifying Approach, Working Paper 00-01, Département d'économique, Université de Sherbrooke.

[15] Feldstein,M.S. (1975), The Income Tax and Charitable Contributions: Part I-Aggregate and Distributional Effects, *National Tax Journal*,28, 81-100.

[16] Fishburn, P.C. and R.D. Willig (1984), Transfer Principles in Income Redistribution, *Journal of Public Economics*, 25, 323-328.

[17] Foster, J., J. Greer and E. Thorbecke (1984), A Class of Decomposable Poverty Measures, *Econometrica*, 52, 761-776.

[18] Foster, J. and A.F. Shorrocks (1988a), Poverty Orderings, *Econometrica*, 56, 173-177.

[19] Foster, J.E. and A.F. Shorrocks (1988b), Poverty Orderings and Welfare Dominance, *Social Choice and Welfare*, 5, 179-198.

[20] Guesnerie, R.(1977), On the Direction of Tax Reform, *Journal of Public Economics*, 7, 179-202.

[21] Härdle, W. (1990), *Applied Nonparametric Regression*, Cambridge University Press, Cambridge.

[22] Kakwani, N. (1980), On a Class of Poverty Measures, *Econometrica*, 48, 437-446.

[23] King, M.A. (1983), Welfare Analysis of Tax Reforms Using Household Data, *Journal of Public Economics*, 21, 183-214.

[24] Kolm, S.C. (1976), Unequal Inequlity: I, *Journal of Economic Theory*, 12, 416-442.

[25] Makdissi, P. and Q. Wodon (2002), Consumption Dominance Curves: Testing for the Impact of Indirect Tax Reforms on Poverty, forthcoming in *Economics Letters*.

[26] Mayshar, J. (1990), On Measures of Excess Burden and their Application, *Journal of Public Economics*, 43, 263-289.

[27] Mayshar, J. and S. Yitzhaki (1995), Dalton Improving Tax Reform, *American Economic Review*, 85, 793-807.

[28] Pen, J. (1971), *Income Distribution: facts, theories, policies*, Preaeger, New York.

[29] Rao, R.C. (1973), *Linear Statistical Inference and Its Applications*, John Wiley and Sons Inc., New York..

[30] Shorrocks, A.F. (1983), Ranking Income Distributions, *Economica*, 50, 3-17.

[31] Shorrocks, A.F. and J.E. Foster (1987), Transfer Sensitive Inequality Measures, *Review of Economic Studies*, 54, 485-497.

[32] Thistle, P. (1993), Negative Moments, Risk Aversion, and Stochastic Dominance, *Journal of Financial Quantitative Analysis*, 28, 301-311.

[33] Watts, H.W. (1968), An Economic Definition of Poverty, in D.P. Moynihan (ed.), *On Understanding Poverty*, Basic Books, New-York.

[34] Wildasin, D. (1984), On Public Good Provision With Distortionary Taxation, *Economic Inquiry*,22, 227-243.

[35] Yitzhaki, S. and J.D. Lewis (1996), Guidelines on Searching for a Dalton-Improving Tax Reform: An Illustration with Data from Indonesia, *World Bank Economic Review*, 10, 541-562.

[36] Yitzhaki, S. and J. Slemrod (1991), Welfare Dominance: An Application to Commodity Taxation, *American Economic Review*, 81, 480-496.

[37] Yitzhaki, S. and W, Thirsk (1990), Welfare Dominance and the Design of Excise Taxation in the Côte d'Ivoire, *Journal of Development Economics*, 33, 1-18.

[38] Zheng, B.(1999), On the Power of Poverty Orderings, *Social Choice and Welfare*, 16, 349-371.

[39] Zoli, C. (1999), Intersecting Generalized Lorenz Curves and the Gini index, *Social Choice and Welfare*, 16, 183-196.

A Appendix 1

A.1 Lemmas

Lemma 9 $\gamma_1(a) < 1$

Proof. Recall that, by definition, we have

$$\int_0^a \overline{x}_l(y)dF(y) = \int_0^a \overline{x}_j(y)dF(y) = 1. \tag{50}$$

Since $\gamma_1(a) \leq \delta_1(y)$ for all $y \in [0, a]$, with strict inequality for some $y \in [0, z^+]$, we also have that

$$\int_0^{z^+} \overline{x}_l(y)dF(y) > \gamma_1(a) \int_0^{z^+} \overline{x}_j(y)dF(y) \tag{51}$$

and

$$\int_{z^+}^a \overline{x}_l(y)dF(y) \geq \gamma_1(a) \int_{z^+}^a \overline{x}_j(y)dF(y). \tag{52}$$

Adding (51) and (52), we find

$$\int_0^a \overline{x}_l(y)dF(y) > \gamma_1(a)\int_0^a \overline{x}_j(y)dF(y) \tag{53}$$

which, in conjunction with (50), implies that

$$\frac{\int_0^a \overline{x}_l(y)dF(y)}{\int_0^a \overline{x}_j(y)dF(y)} = 1 > \gamma_1(a).$$

■

Lemma 10 $\gamma_2(a) \leq 1$.

Proof. Note that, by definition, $\delta^2(a) = 1$ since $\overline{CD}_j^2(a) = \overline{CD}_l^2(a) = 1$. Since $\gamma_2(a) = \inf\{\delta^s(z)\,|z \in [0,a]\}$, it must be that $\gamma_2(a) \leq 1$. ■

Lemma 11 *If a tax reform is s-order socially efficient up to some z^+, then it is also socially efficient up to z^+ at the $s+1$ order. Furthermore, if a tax reform is s-order socially efficient up to some z^+, and if there is strict s-order dominance over at least part of a bottom range of $z \in [0,w]$, then $\gamma_{s+1}(z^+) \geq \gamma_s(z^+)$, with strict inequality if $\gamma_s(z^+) < z^{++}$.*

Proof. We have that $\overline{CD}_l^s(z) - \gamma_s(z^+)\overline{CD}_j^s(z) \geq 0, \forall z \in [0,z^+]$. For $(s+1)$-order social efficiency, we consider the sign of $\overline{CD}_l^{s+1}(z) - \gamma_s(z^+)\overline{CD}_j^{s+1}(z)$. Again, this equals:

$$\overline{CD}_l^{s+1}(z) - \gamma_s(z^+)\overline{CD}_j^{s+1}(z) = \int_0^z \left(\overline{CD}_l^s(y) - \gamma_s(z^+)\overline{CD}_j^s(y)\right)dy, \tag{54}$$

which, by assumption, is clearly positive for all $z \in [0,z^+]$. This proves the first part of the lemma. If there is strict dominance over a bottom range of $z \in [0,w]$, $w \leq z^+$, then (54) can be split into

$$\int_0^w \left(\overline{CD}_l^s(y) - \gamma_s(z^+)\overline{CD}_j^s(y)\right)dy + \int_w^{z^+} \left(\overline{CD}_l^s(y) - \gamma_s(z^+)\overline{CD}_j^s(y)\right)dy \tag{55}$$

The first term of (55) is strictly positive, while the second is non-negative. Hence, $\overline{CD}_l^{s+1}(z) - \gamma_s(z^+)\overline{CD}_j^{s+1}(z) > 0, \quad \forall z \in [0,z^+]$. Hence, it must be that $\gamma_{s+1}(z^+) \geq \gamma_s(z^+)$, with strict inequality if $\gamma_s(z^+) < z^{++}$. ■

Lemma 12 *Let a tax reform be s-order socially efficient up to some* $z_s(\gamma^+)$*, for some given* γ^+*, with strict dominance over at least part of* $[0, z_s(\gamma^+)]$*. Then,* $z_{s+1}(\gamma^+) \geq z_s(\gamma^+)$*, with strict inequality if* $z_s(\gamma^+) < z^{++}$*.*

Proof. Denote the lower bound of the interval of strict s-order dominance by w, with $w < z^+$. For $(s+1)$-order social efficiency, we consider the sign of $\overline{CD}_l^{s+1}(z) - \gamma^+ \overline{CD}_j^{s+1}(z)$. This equals:

$$\overline{CD}_l^{s+1}(z) - \gamma^+ \overline{CD}_j^{s+1}(z) = \int_0^z \left(\overline{CD}_l^s(y) - \gamma^+ \overline{CD}_j^s(y) \right) dy. \tag{56}$$

Hence, we have that

$$\begin{array}{c} \overline{CD}_l^{s+1}(z) - \gamma^+ \overline{CD}_j^{s+1}(z) \geq 0 \text{ if } z \in [0, w] \\ \overline{CD}_l^{s+1}(z) - \gamma^+ \overline{CD}_j^{s+1}(z) > 0 \text{ if } z \in]w, z_s(\gamma^+)]. \end{array} \tag{57}$$

Recall that
$z_{s+1}(\gamma^+) = z_{s+1}(\gamma^+) = \sup \left\{ z \left| \overline{CD}_l^{s+1}(z) - \gamma^+ \overline{CD}_j^{s+1}(z), y \in [0, z], z \leq z^{++} \right. \right\}$.
From (57), note that $\overline{CD}_l^{s+1}(z_s(\gamma^+)) - \gamma^+ \overline{CD}_j^{s+1}(z_s(\gamma^+)) > 0$. Hence, it must be that $z_{s+1}(\gamma^+) > z_s(\gamma^+)$ if $z_s(\gamma^+)$ is not constrained by the upper bound z^{++}. ■

Lemma 13 $\gamma_0(z^+) \leq \gamma_1(z^+)$*, with strict inequality if there is heterogeneity in the ratio of Engel curves for goods* x_l *and* x_j *at each* $y \in [0, z^+]$ *and if* $\gamma_1(z^+) < \gamma^{++}$*.*

Proof. $\gamma_1(z^+)$ is such that $\forall y \in [0, z^+]$.

First, let $\gamma_1(z^+) = \gamma^{++}$. This implies that $\delta^1(y) = \gamma^{++}$, that $\overline{x}_j(y) = 0$ and thus that $\overline{x}_j(y, \theta) = 0$ and $\delta^0(y, \theta) = \gamma^{++}$ for all θ for which $dF(\theta \,|y) > 0$, and for all $y \in [0, z^+]$. Hence, $\gamma_0(z^+) = \gamma^{++} \leq \gamma_1(z^+)$.

Second, let $\gamma_1(z^+) < \gamma^{++}$. Then, there exists a $y^* \in [0, z^+]$ such that $\gamma_1(z^+) = \delta^1(y^*) = \overline{x}_l(y^*)/\overline{x}_j(y^*)$. Let Ω_1 and Ω_2 be two exclusive and exhaustive subsets of preferences at y^* with values of Engel curve ratios $\overline{x}_l(y^*, \theta)/\overline{x}_j(y^*, \theta)$ set respectively to $\rho_1(y^*)$ and $\rho_2(y^*)$ (a proof with an arbitrarily greater number of subsets follows along the same lines). Then, by definition,

$$\begin{aligned}
\delta^1(y^*) &= \frac{\overline{x}_l(y^*)}{\overline{x}_j(y^*)} \\
&= \frac{\int_{\Omega_1} \overline{x}_l(y^*,\theta) dF(\theta\,|y^*) + \int_{\Omega_2} \overline{x}_l(y^*,\theta) dF(\theta\,|y^*)}{\overline{x}_j(y^*)} \\
&= \rho_1(y^*)s + \rho_2(y^*)(1-s)
\end{aligned}$$

where

$$s = \frac{\int_{\Omega_1} dF(\theta\,|y^*)}{\overline{x}_j(y^*)}. \tag{58}$$

Hence, in the presence of heterogeneity at y^* (which implies that $\rho_1(y^*) \neq \rho_2(y^*)$ and $s > 0$), we find by (58) that $\delta^1(y^*) > \min(\rho_1(y^*), \rho_2(y^*))$. Since $\gamma_0(z^+) \leq \min((\rho_1(y^*), \rho_2(y^*))$ and $\gamma_1(z^+) = \delta^1(y^*)$, we have that $\gamma_0(z^+) < \gamma_1(z^+)$. ■

Lemma 14 $z_0(\gamma^+) \leq z_1(\gamma^+)$, *with strict inequality if there is heterogeneity in the ratio of Engel curves for goods* x_l *and* x_j *at* $y = z_1(\gamma^+)$ *and if* $z_1(\gamma^+) < z^{++}$.

Proof. Denote by $\rho_1(y)$ and $\rho_2(y)$ two existing values of ratios of Engel curves at y. By the proof of the preceding Lemma 13, we have that $\delta^1(y) \geq \min(\rho_1(y),\rho_2(y))$. Hence, by equation (33),

$$\begin{aligned}
z_0(\gamma^+) &= \sup\left\{ z \,\middle|\, \begin{array}{c} \delta^0(y,\theta) \geq \gamma^+, y \in [0,z], \theta \in \Omega, \\ dF(y) > 0,\ dF(\theta|y) > 0, z \leq z^{++} \end{array} \right\} \\
&\leq \sup\left\{ z \,\middle|\min(\rho_1(y), \rho_2(y)) \geq \gamma^+, y \in [0,z], dF(y) > 0, z \leq z^{++} \right\} \\
&\leq \sup\left\{ z \,\middle|\delta^1(y) \geq \gamma^+, y \in [0,z], z \leq z^{++} \right\} \\
&= z_1(\gamma^+).
\end{aligned}$$

Now, suppose that $z_1(\gamma^+) < z^{++}$. We then have that $\delta^1(z_1(\gamma^+)) = \gamma^+$, and if there is heterogeneity at $z_1(\gamma^+)$, that $\gamma^+ = \delta^1(z_1(\gamma^+)) > \min(\rho_1(y), \rho_2(y))$. Thus, by (33), $z_0(\gamma^+) \leq z_1(\gamma^+)$ since

$$\min_{\theta\in\Omega, dF(\theta|z_1(\gamma^+))>0} \delta^0(z_1(\gamma^+),\theta) \leq \min(\rho_1(y), \rho_2(y)) < \gamma^+.$$

■

A.2 Proofs of Theorems

Proof. Proof of Theorem 1.

Imagine two distributions A (before a tax reform) and B (after a tax reform). Duclos and Makdissi (2000) show that a necessary and sufficient condition for poverty to decrease when moving from A to B, for all $P(z) \in \Pi^s(z)$, for all $z \in [0, z^+]$, and for any given $s \in \{1, 2, 3, ...\}$, is that

$$D_A^s(y) \geq D_B^s(y) \ \ \forall y \leq z^+. \tag{59}$$

Note that the continuity assumption $p^{(t)}(z, z) = 0$ for all $t \in \{1, 2, ..., s-2\}$ is important for ordering distributions at dominance order 3 and higher. In the context of a marginal tax reform, this necessary and sufficient condition becomes

$$dD^s(y) \leq 0 \ \forall y \leq z^+. \tag{60}$$

We have

$$dD^s(y) = \frac{\partial D^s(y)}{\partial t_l} dt_l + \frac{\partial D^s(y)}{\partial t_j} dt_j. \tag{61}$$

Using (8), (61) may be rewritten as

$$\begin{aligned} dD^s(y) &= \left[\frac{\partial D^s(y)}{\partial t_l} - \gamma \frac{X_l}{X_j} \frac{\partial D^s(y)}{\partial t_j}\right] dt_l & (62) \\ &= \left[\frac{1}{X_l} \frac{\partial D^s(y)}{\partial t_l} - \gamma \frac{1}{X_j} \frac{\partial D^s(y)}{\partial t_j}\right] X_l dt_l. & (63) \end{aligned}$$

Using equations (17) and (18), we obtain

$$dD^s(y) = \left[\overline{CD}_l^s(y) - \gamma \overline{CD}_j^s(y)\right] X_l dt_l. \tag{64}$$

Considering that X_l is positive and that dt_l is negative, condition (60) is then equivalent to

$$\overline{CD}_l^s(y) - \gamma \overline{CD}_k^s(y) \geq 0, \ \forall y \in [0, z^+]. \tag{65}$$

■

Proof. Proof of Theorem 2.

Consider again two distributions A (before a tax reform) and B (after a tax reform). It is well known (see Thistle (1993) and Duclos and Makdissi (2000)) that a sufficient condition for welfare to increase, for all $U \in \Omega^s$ and for any given $s \in \{1, 2, 3, ...\}$, is that

$$D_A^s(y) \geq D_B^s(y) \;\; \forall y \leq a, \tag{66}$$

as well as for $s \geq 3$ that

$$D_A^i(a) \geq D_B^i(a) \;\; \forall i \in \{2, ..., s-1\}. \tag{67}$$

This test simplifies if you use a result of Duclos and Makdissi (2000). To see how, first define a binary relation $>_s^*$: ■

Definition 15 *$D_A(y) >_s^* D_B(y)$ if and only if $D_A^i(y) - D_B^i(y) > 0$ for the smallest $i \leq s$ such that $D_A^i(y) \neq D_B^i(y)$.*

We then obtain the following lemma.

Lemma 16 (Duclos and Makdissi (2000)) *$D_A(a) >_s^* D_B(a)$ if $D_A^s(y) - D_B^s(y) \geq 0 \; \forall x \in \Re_+$.*

This establishes a link between the s-order dominance conditions over $\Re_+$ and the dominance conditions at the limit a of the distribution at some lower order. Note, moreover, that when $D_A(a) >_s^* D_B(a)$, it must be that $\lim_{y\to\infty}[D_A^i(y) - D_B^i(y)] \geq 0$ for all $i = 1, 2, ...$ (see Duclos and Makdissi (2000)). We thus have:

Corollary 17 *If $D_A^s(y) - D_B^s(y) \geq 0 \; \forall y \in \Re_+$, then $\lim_{y\to\infty}[D_A^i(y) - D_B^i(y)] \geq 0$ for all $i \leq s$.*

Using this last result, we can state that a sufficient condition for welfare to increase, for all $U \in \Omega^s$ and for any given $s \in \{1, 2, 3, ...\}$, is that

$$D_A^s(y) \geq D_B^s(y) \;\; \forall y \in \Re_+. \tag{68}$$

In the context of a marginal tax reform, this necessary and sufficient condition becomes

$$dD^s(y) \leq 0 \; \forall y \in \Re_+. \tag{69}$$

We can the follow the proof of Theorem 1 to show that this is equivalent to testing whether

$$\overline{CD}_l^s(y) - \gamma \overline{CD}_k^s(y) \geq 0 \ \forall y \in \Re_+. \tag{70}$$

Note from Lemma 10 that we cannot have Dalton efficiency if $\gamma > 1$. Hence, by Corollary 17, we cannot have higher-order social efficiency either.

Proof. Proof of Theorem 5.

$\widehat{CD}_k^s(z)$ is a consistent estimator of $CD_k^s(z)$ by the existence of the first population moment of $x_k(y,\theta)(z-y)_+^{s-2}$ and the law of large numbers. $\widehat{CD}_k^s(z)$ is $N^{0.5}$ consistent and asymptotically normal by the existence of the second population moment and the central limit theorem, with asymptotic variance given by 35 by simple calculation. ■

Proof. Proof of Theorem 6.

Note first that $E\left[\widehat{CD}_k^1(z)\right] = \int \kappa_h(z-y)\, x_k(y) f(y) dy$. Denoting $t = h^{-1}(z-y)$ and expanding around $t_0 = 0$, for small h this is approximately equal to

$$\begin{aligned} E\left[\widehat{CD}_k^1(z)\right] &\simeq \int \kappa(t)\left[CD_k^1(z) - thCD_k^{1'}(z) + 0.5t^2h^2CD_k^{1''}(z)\right]dt \\ &= CD_k^1(z) + 0.5h^2CD_k^{1''}(z)\, c_\kappa \end{aligned}$$

since $\int \kappa(u)\,du = 1$, $\int u\kappa(u)\,du = 0$, and $\int u^2\kappa(u)\,du = c_\kappa$. Hence, the bias $E\left[\widehat{CD}_k^1(z)\right] - CD_k^1(z)$ is given by $0.5h^2CD_k^{1''}(z)\,c_\kappa$.

By (39), note that $\widehat{CD}_k^1(z)$ is a sum of independent iid variables to which we may apply the central limit theorem and show asymptotic normality. We also have that

$$\begin{aligned} & N \operatorname{var}\left(\widehat{CD}_k^1(z)\right) \\ = \;& \operatorname{var}\left(\kappa_h(z-y)\, x_k(y,\theta)\right) = E\left[\kappa_h(z-y)^2\left(x_k(y,\theta)\right)^2\right] - E\left[\widehat{CD}_k^1(z)\right]^2 \\ = \;& \int_y\int_\Omega \kappa_h(z-y)^2\left(x_k(y,\theta)\right)^2 dF(\theta|y)dF(y) - E\left[\widehat{CD}_k^1(z)\right]^2 \\ = \;& \int_u\int_\Omega h^{-2}\kappa(u)^2\left(x_k(z-uh,\theta)\right)^2 dF(\theta|z-uh)dF(z-uh) - E\left[\widehat{CD}_k^1(z)\right]^2 \end{aligned} \tag{71}$$

where the last expression is obtained by substituting u for $h^{-1}(z-y)$. For small h, (71) is approximately equal to

$$
\begin{aligned}
& N\text{var}\left(\widehat{CD}_k^1(z)\right) \\
\cong\ & \int_u \int_\Omega h^{-1}\kappa(u)^2 (x_k(z,\theta))^2 dF(\theta|z) f(z) du - E\left[\widehat{CD}_k^1(z)\right]^2 \\
=\ & h^{-1} f(z) \int_\Omega (x_k(z,\theta))^2 dF(\theta|z) \int_u \kappa(u)^2 du - E\left[\widehat{CD}_k^1(z)\right]^2 \qquad (72) \\
\cong\ & h^{-1} f(z) \int_\Omega (x_k(z,\theta))^2 dF(\theta|z) \int_u \kappa(u)^2 du \qquad (73) \\
=\ & h^{-1} f(z) c_k(z) \int \kappa(u)^2 du. \qquad (74)
\end{aligned}
$$

Hence,

$$
\begin{aligned}
& \lim_{N\to\infty} Nh\, var\left(\widehat{CD}_k^1(z) - CD_k^1(z) - h^2 B_k\right) \\
=\ & f(z) c_k(z) \int \kappa(u)^2 du = V_k(z),
\end{aligned}
$$

which concludes the proof. ■

Proof. Proofs of Theorems 7 and 8.

The proofs follow the lines of the proof of Theorem 3 in Davidson and Duclos (2000), and we will therefore simply outline them. The proof for $\widehat{\gamma}_s(z^+)$ is similar to that of $\widehat{z}_s(\gamma^+)$, and we thus focus on the proof for $\widehat{z}_s(\gamma^+)$.

Under the assumption that $z_s(\gamma^+)$ exists in the population (and is less than z^{++}), $\widehat{z}_s(\gamma^+)$ is a consistent estimator of $z_s(\gamma^+)$ for all $s = 1, 2, ...$ We can also show that

$$
\widehat{d}_s(\widehat{z}_s(\gamma^+), \gamma^+) - \widehat{d}_s(z_s(\gamma^+), \gamma^+) - \left(d_s(\widehat{z}_s(\gamma^+), \gamma^+) - d_s(z_s(\gamma^+), \gamma^+)\right) \\
= o_s
$$

with $o_s = o(N^{-0.5})$ if $s \geq 2$ and $o_1 = o(N^{-0.4})$, since $\widehat{z}_s(\gamma^+) - z_s(\gamma^+) = o(1)$ for all s. Since, by definition, $\widehat{d}_s(\widehat{z}_s(\gamma^+), \gamma^+) = d_s(z_s(\gamma^+), \gamma^+) = 0$, it thus follows that

$$
\widehat{d}_s(z_s(\gamma^+), \gamma^+) = d_s(\widehat{z}_s(\gamma^+), \gamma^+) + o_s. \qquad (75)
$$

By a Taylor expansion of $d_s(\widehat{z}_s(\gamma^+), \gamma^+)$ around $z_s(\gamma^+)$, we have that:

$$
d_s(\widehat{z}_s(\gamma^+), \gamma^+) = d_s^z(\widetilde{z}, \gamma^+)\left(\widehat{z}_s(\gamma^+) - z_s(\gamma^+)\right) \qquad (76)
$$

for some $\widetilde{z}$ such that $|\widetilde{z} - z_s(\gamma^+)| < |\widehat{z}_s(\gamma^+) - z_s(\gamma^+)|$. We can thus write from (75) and (76):

$$\widehat{z}_s(\gamma^+) - z_s(\gamma^+) = \frac{d_s(\widehat{z}_s(\gamma^+), \gamma^+)}{d_s^z(\widetilde{z}, \gamma^+)} + o_s. \tag{77}$$

We can apply the central limit theorem to $N^{0.5} d_s(\widehat{z}_s(\gamma^+), \gamma^+)$ for $s \geq 2$ and to $(Nh)^{0.5}\, d_1(\widehat{z}_s(\gamma^+), \gamma^+)$ for $s = 1$ since by (77) they are sums of iid variables, since the second population moment of $(z_s(\gamma^+) - y)_+^{s-2}\,(x_j - \gamma^+ x_l)$ has been assumed to exist (for $s \geq 2$), and since the conditions of Theorem 6 were assumed to hold for $s = 1$. Hence, $N^{0.4}\,(\widehat{z}_1(\gamma^+) - z_1(\gamma^+)) = O(1)$ and $N^{0.5}\,(\widehat{z}_s(\gamma^+) - z_s(\gamma^+)) = O(1)$ for $s \geq 2$ and are asymptotically normal. Because $\widehat{z}_s(\gamma^+) \to z_s(\gamma^+)$ as $N \to \infty$, $\widetilde{z} \to z_s(\gamma^+)$, we have that $d_s^z(\widetilde{z}, \gamma^+)$ will converge to $d_s^z(z_s(\gamma^+), \gamma^+) \neq 0$ for $s = 1, 2, ...$ The asymptotic variance then follows by computation. ■

Figure 1: CD curves and critical social efficiency parameters

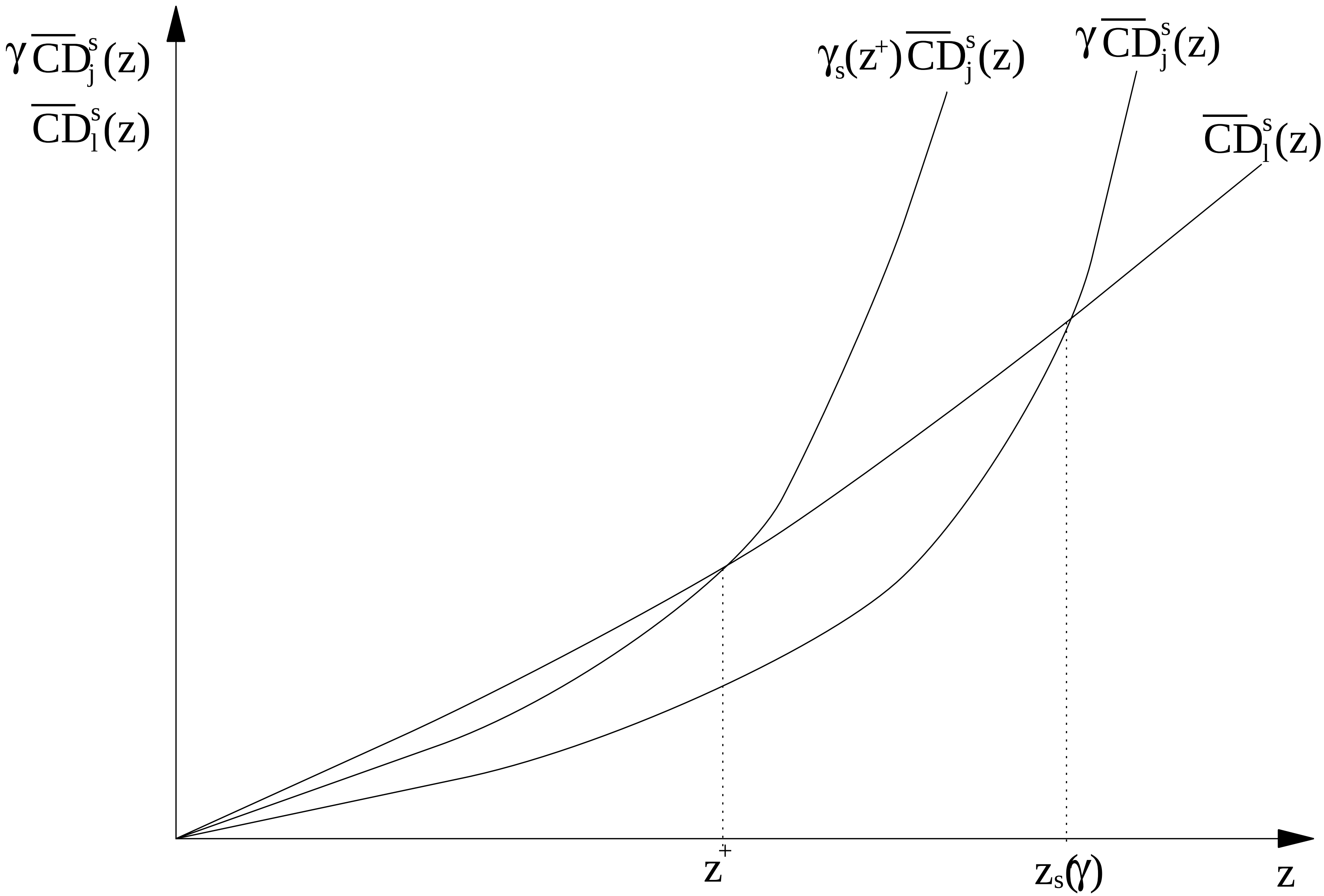


Figure 2: Distributive benefit and economic efficiency

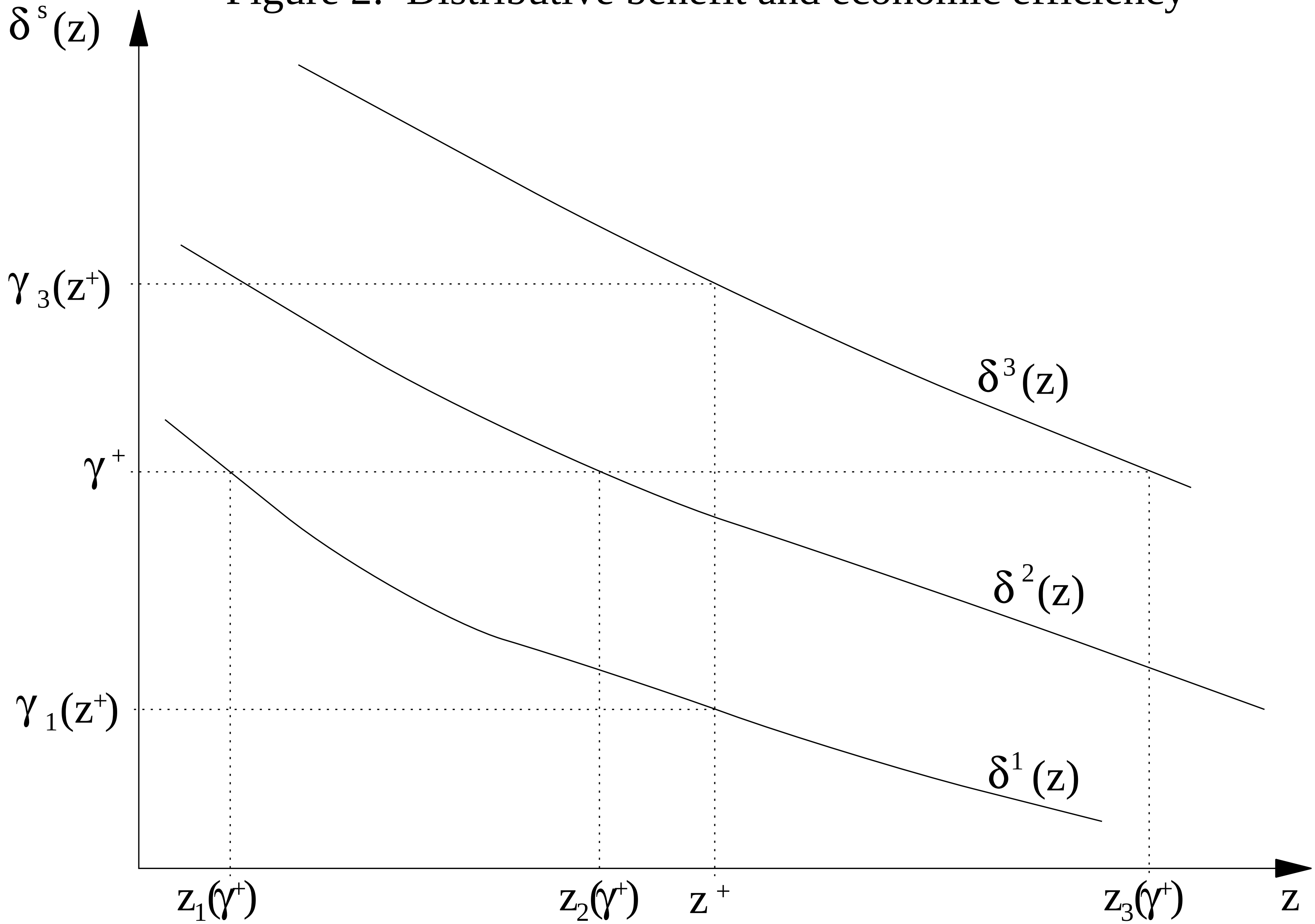
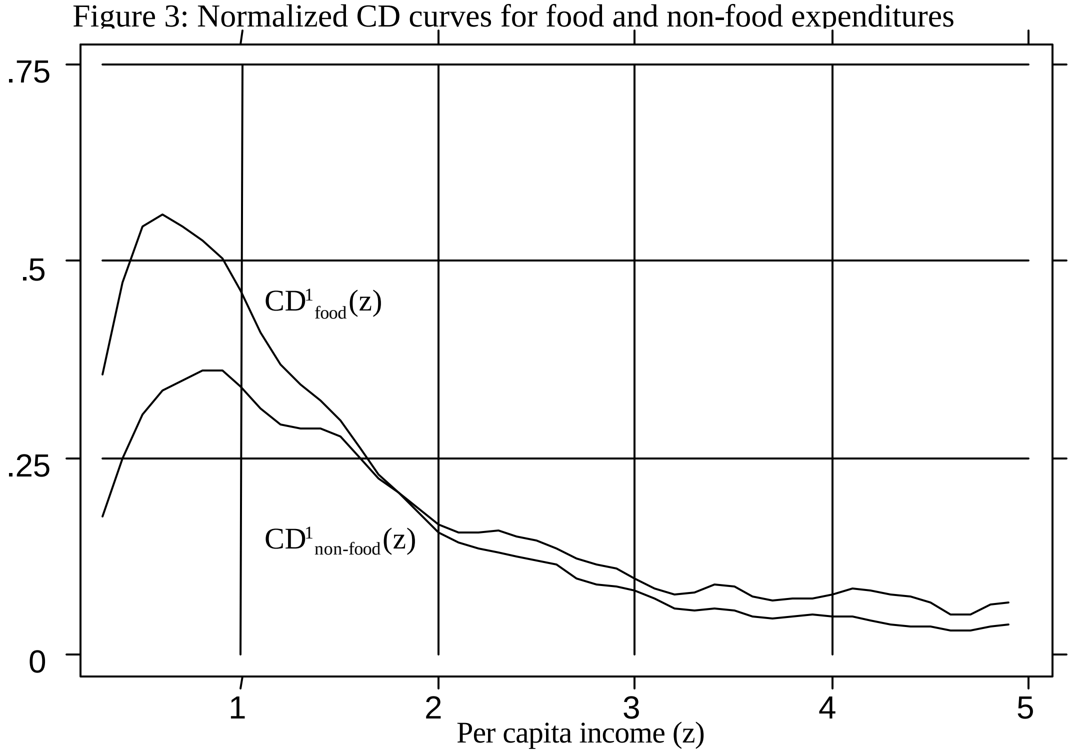
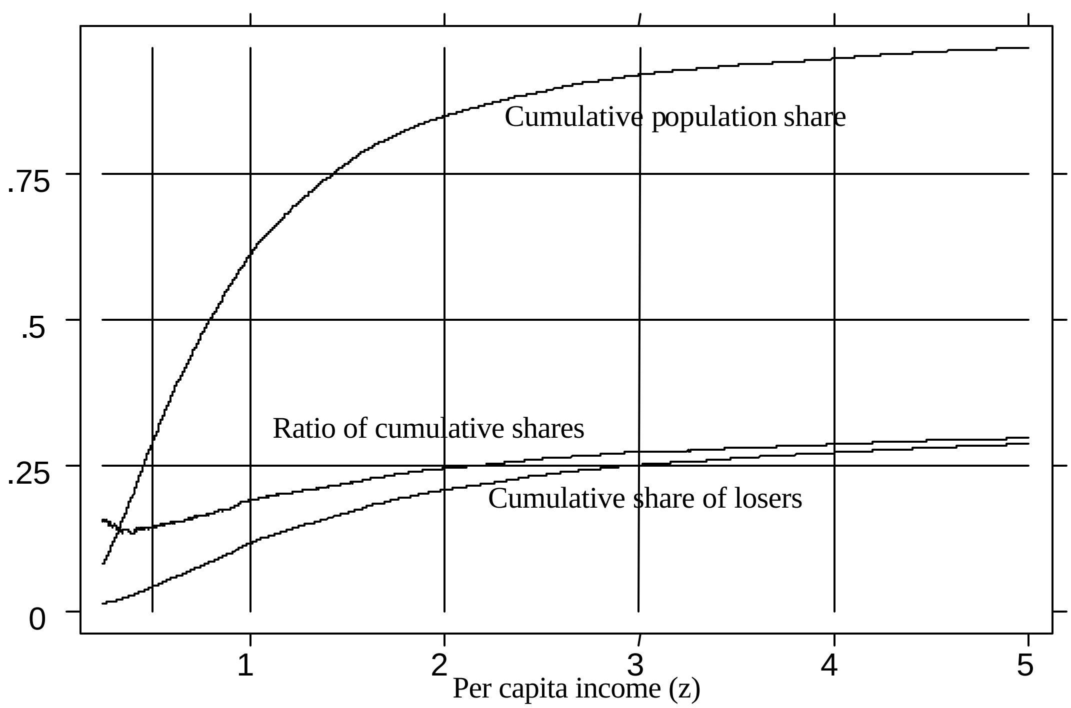


Figure 3: Normalized CD curves for food and non-food expenditures



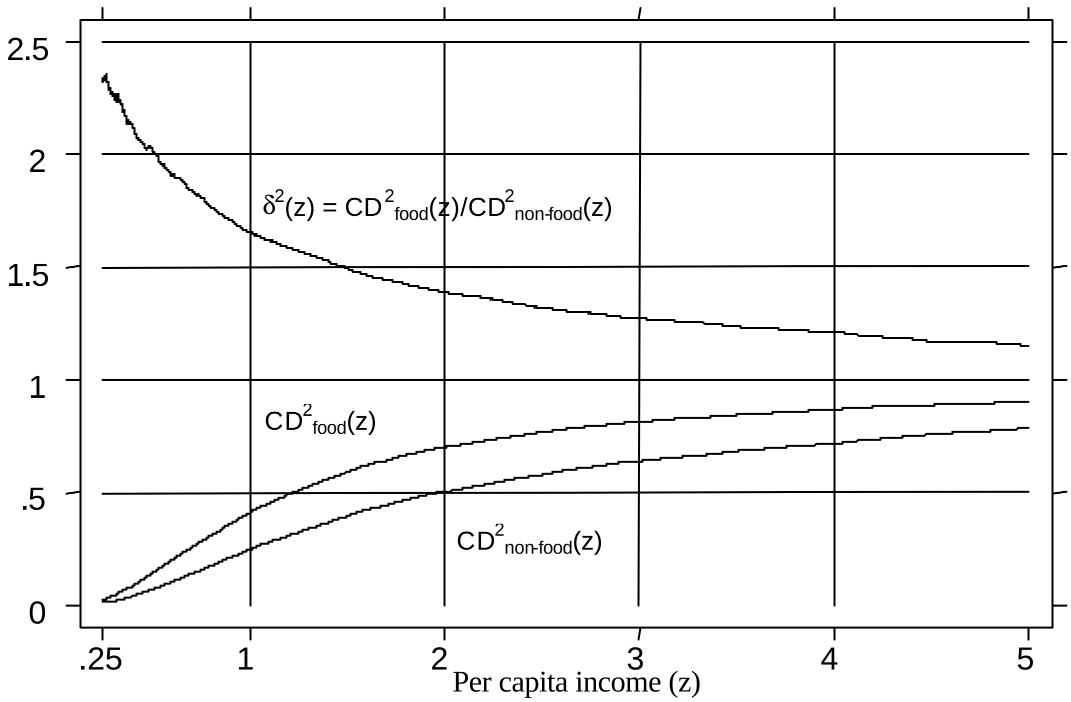
Source: Authors' estimation using 1996 ENIGH.

Figure 4: Share of losers from food/non-food tax reform



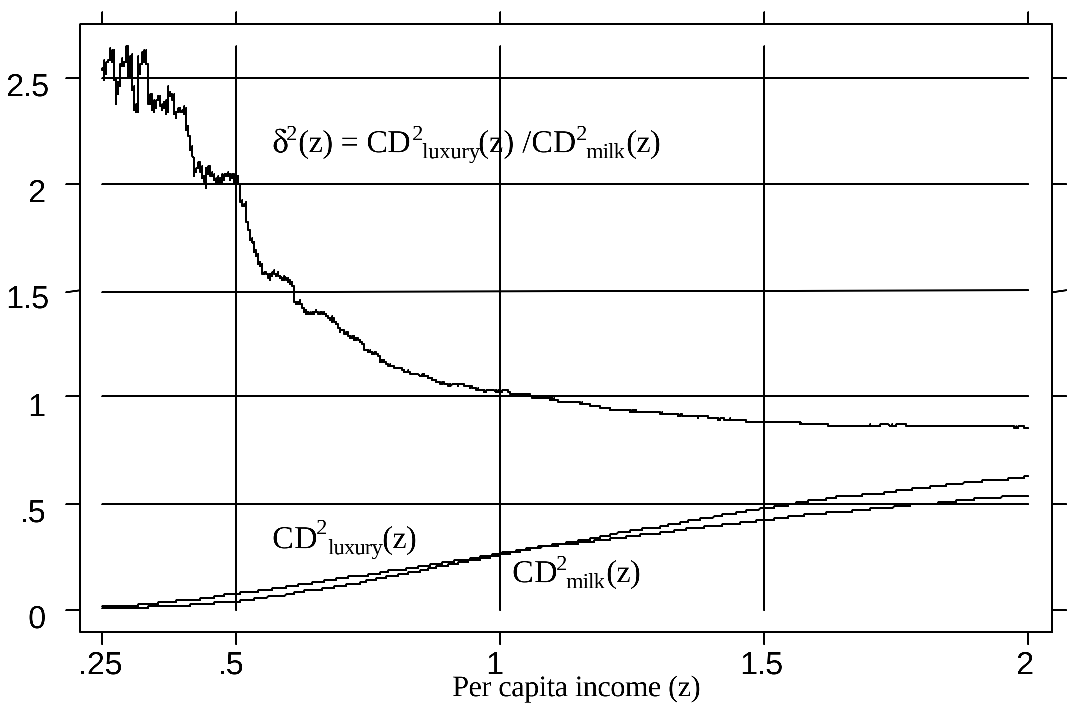
Source: Authors' estimation using 1996 ENIGH.

Figure 5: CD curves for food and non-food expenditures, s=2



Source: Authors' estimation using 1996 ENIGH.

Figure 6: CD curves for mixed bundle and pasteurized milk, s=2



Source: Authors' estimation using 1996 ENIGH.

Table 1: Indirect taxation for food versus non-food expenditures, Mexico 1996

	Value of second-order CD curves at various poverty lines z		
	$\overline{CD}^2_{food}(z)$	$\overline{CD}^2_{non-food}(z)$	$\overline{CD}^2_{food}(z)-\overline{CD}^2_{non-food}(z)$
z=0.5	0.141	0.070	0.071
	(0.004)	(0.002)	(0.002)
z=1	0.410	0.247	0.163
	(0.007)	(0.006)	(0.004)
z=2	0.700	0.504	0.196
	(0.007)	(0.010)	(0.006)
	Critical efficiency ratios $g_s(z^+)$ for different maximum poverty lines z^+ and for different orders of dominance s		
	z^+=0.5 (28.5% of population covered)	z^+=1 (60.9% of population covered)	z^+=2 (85.0% of population covered)
$g_1(z^+)$	1.782	1.354	0.947
	(0.032)	(0.027)	(0.029)
$g_2(z^+)$	2.021	1.657	1.390
	(0.041)	(0.025)	(0.018)
$g_3(z^+)$	2.140	1.822	1.551
	(0.049)	(0.028)	(0.020)
	Critical poverty lines $z_s(g)$ for different ratios of economic efficiency costs g and for different orders of dominance s (*)		
	g=0.5	g=1.0	g=1.5
$z_1(g)$	4.272	1.793	0.752
	(0.127)	(0.099)	(0.030)
$z_2(g)$	-	-	1.483
			(0.156)
$z_3(g)$	-	-	2.347
			(0.145)

Source: Authors' estimation using 1996 ENIGH. Sample size is 14022 observations. Standard errors in parenthesis. (*) Only poverty lines higher than 0.25 are considered.

Table 2: Indirect taxation for mixed food bundle and pasteurized milk, Mexico 1996

	Value of second-order CD curves at various poverty lines z		
	$\overline{CD}^2_{mixed}(z)$	$\overline{CD}^2_{milk}(z)$	$\overline{CD}^2_{mixed}(z) - \overline{CD}^2_{milk}(z)$
z=0.5	0.075	0.037	0.038
	(0.005)	(0.003)	(0.005)
z=1	0.263	0.256	0.007
	(0.009)	(0.008)	(0.010)
z=2	0.537	0.627	-0.090
	(0.012)	(0.011)	(0.012)
	Critical efficiency ratios $g_s(z^+)$ for different maximum poverty lines z^+ and for different orders of dominance s		
	z^+=0.5 (28.5% of population covered)	z^+=1 (60.9% of population covered)	z^+=2 (85.0% of population covered)
$g_1(z^+)$	1.222	0.800	0.705
	(0.080)	(0.043)	(0.072)
$g_2(z^+)$	1.997	1.028	0.856
	(0.181)	(0.039)	(0.019)
$g_3(z^+)$	2.214	1.256	0.946
	(0.250)	(0.059)	(0.025)
	Critical poverty lines $z_s(g)$ for different ratios g of economic efficiency costs and for different orders of dominance s (*)		
	g=0.5	g=1.0	g=1.5
$z_1(g)$	11.390	0.600	0.421
	(0.065)	(0.033)	(0.028)
$z_2(g)$	-	1.062	0.613
		(0.212)	(0.083)
$z_3(g)$	-	1.623	0.795
		(0.140)	(0.052)

Source: Authors' estimation using 1996 ENIGH. Sample size is 14022 observations. Standard errors in parenthesis. (*) Only poverty lines higher than 0.25 are considered.