The Lamé Class of Lorenz Curves

José María Sarabia[1], Vanesa Jordá, Carmen Trueba

Department of Economics, University of Cantabria, 39005 Santander, Spain

Abstract

In this paper, the class of Lamé Lorenz curves is studied. This family has the advantage of modeling inequality with a single parameter. The family has a double motivation: it can be obtain from an economic model and from simple transformations of classical Lorenz curves. The underlying cumulative distribution functions have a simple closed form, and correspond to the Singh-Maddala and Dagum distributions, which are well known in the economic literature. The Lorenz order is studied and several inequality and polarization measures are obtained, including Gini, Donaldson-Weymark-Kakwani, Pietra and Wolfson indices. Some extensions of the Lamé family are obtained. Fitting and estimation methods under two different data configuration are proposed. Empirical applications with real data are given. Finally, some relationships with other curves are included.

Key Words: Lorenz curves, Gini, Donaldson-Weymark-Kakwani, Pietra and Wolfson indices, limited information.

1   Introduction

The Lorenz curve (LC) is an important and convenient statistical instrument used in the analysis of economical data. The setting of obtaining an adequate functional form for this curve, continues being an active research field. Different parametric models for LCs, have been proposed by Kakwani and Podder (1976), Rasche et al. (1980), Aggarwal (1984), Aggarwal and Singh (1984),

---

[1]Corresponding author. Tel.: 34 942201635; Fax: 34 942201603. E-mail address: sarabiaj@unican.es (JM Sarabia).

Arnold (1986), Pakes (1986), Arnold et al. (1987), Villaseñor and Arnold (1989), Basmann et al. (1990), Ortega et al. (1991), Chotikapanich (1993), Holm (1993), Ryu and Slottje (1996), Sarabia (1997), Sarabia, Castillo and Slottje (1999, 2001), Ogwang and Rao (1996, 2000), Sarabia and Pascual (2002), Sarabia et al (2005, 2010), Rhode (2009) and Helene (2010).

In a recent paper, Henle et al. (2008) introduces a family of Lorenz curves, the so called Lamé family, which is characterized by a single inequality parameter. This family presents several practical and theoretical advantages. The authors obtain the family from an economic model (see Section 3). On the other hand, the idea of describing the inequality by a single parameter is quite interesting from an economic point of view, and has been supported by Slottje (2010). The family of Lamé Lorenz curves is defined by two simple functional forms, which depend on one single parameter. One of the main advantages of this family, is that we can obtain estimates of the Gini index another inequality measures very similar to those obtained with other more complicated functional forms.

In this paper, the class of Lamé Lorenz curves is studied. This family has the advantage of modeling inequality with a single parameter, and many of the most important statistical and economic measures for studying the inequality can be obtained in a closed form. The family has a double motivation: it can be obtain from an economic model and from a simple transformation of the potential and classical Pareto Lorenz curves. The underlying cumulative distribution functions have a simple closed form, and correspond to the Singh-Maddala and Dagum distributions, which are very popular in the economic literature. The Lorenz order is studied and several inequality and polarization measures are obtained, including Gini, Donaldson-Weymark-Kakwani, Pietra and Wolfson indices. Two direct and simple extensions of the Lamé family are obtained. Fitting and estimation methods under two different data configuration are proposed. Two empirical applications with real data are presented. Finally, some relationships with other curves are included.

The rest of this paper is organized as follows. In Section 2 the basic results, which will be used along the following paper, will be introduced. The family of Lamé Lorenz curves is introduced in Section 3 and the underlying cumulative distribution function is obtained in Section 4. The Lorenz ordering and the inequality and polarization measures are studied in Section 5. Some extensions of the Lamé Lorenz curves are discussed in section 6. Estimation methods are proposed in Section 7. Section 8 includes two empirical

applications. Finally, some relationships with other curves are included in Section 9.

2 Previous Results

In this paper we consider the Lorenz curve as defined by Gastwirth (1971). Let $\mathcal{L}$ be the class of all non-negative random variables with positive finite expectation μ. For a random variable X in $\mathcal{L}$ with distribution function F_X we define its inverse distribution function F_X^{-1} by

$$F_X^{-1}(y) = \sup\{x : F_X(x) \leq y\}.$$

Thus, the Lorenz curve associated with X is defined by,

$$L_X(p) = \left[\int_0^p F_X^{-1}(y)dy \right] / \left[\int_0^1 F_X^{-1}(y)dy \right], \quad 0 \leq p \leq 1. \qquad (2.1)$$

Note that $\mu_X = \int_0^1 F_X^{-1}(y)dy$ is the expectation of the random variable X. A detailed discussion about these concepts can be found in Arnold (1987), Johnson, Kotz and Balakrishnan (1994, Chapter 12) and Sarabia (2008a).

A characterization of a Lorenz curve attributed to Gaffney and Anstis by Pakes (1986) is given by the following theorem.

Theorem 2.1 *Suppose that $L(p)$ is defined and continuous on $[0,1]$ with second derivative $L''(p)$. The function $L(p)$ is a Lorenz curve if and only if*

$$L(0) = 0, \quad L(1) = 1, \quad L'(0^+) \geq 0, \quad L''(p) \geq 0 \; for \, p \in (0,1). \qquad (2.2)$$

Note that $L'(0^+)$ means the value of first derivative of function $L(p)$ for $p = 0$ to the right.

The following lemma considers the composition of Lorenz curves.

Lemma 2.1 *Let $L_i(p)$, $i = 1, 2$ two genuine Lorenz curves. Then, the functional form*

$$L_{12}(p) = L_1[L_2(p)], \quad 0 \leq p \leq 1,$$

defines a new genuine Lorenz curve.

Proof: The proof is direct checking conditions (2.2) in Theorem 2.1. ∎

Using (2.1) it is clear that $F_X^{-1}(y) = \mu_X L_X'(y)$. In relation to the probability density function (PDF) $f(x)$ associated with a Lorenz curve $L(p)$ we have the following (Arnold, 1987),

Theorem 2.2 *If the second derivative $L''(p)$ exists and is positive everywhere in an interval (x_1, x_2), the corresponding cumulative distribution function F has a finite positive probability density function in the interval $(\mu L'(x_1^+), \mu L'(x_2^-))$, which is given by*

$$f(x) = \frac{1}{\mu L'' [F(x)]}.$$

3 The Family

The Lamé class of Lorenz curves is given by the curves,

$$L_1(p; a) = [1 - (1 - p)^a]^{1/a}, \;\; 0 \leq p \leq 1, \tag{3.1}$$

and

$$L_2(p; a) = 1 - (1 - p^a)^{1/a}, \;\; 0 \leq p \leq 1. \tag{3.2}$$

It can be proved using Theorem 2.1 that (3.1) is a genuine Lorenz curve if $0 < a \leq 1$ and (3.2) is a genuine Lorenz curve if $a \geq 1$. Note that for both the curves the case $a = 1$ corresponds to the egalitarian line.

3.1 Motivation

The family has a double motivation: it can be obtain from an economic model and from simple transformations of classical Lorenz curves.

In relation with the economic motivation, Henle, Horton and Jakus (2008) obtained the curves (3.1) and (3.2) making use two different economic theories, the so called trickle-up and trickle-down effects. The first approach supposes that an increase in income of the lower-middle class would be more advantage for the economy since they spend their wealth faster than the upper class. In contrast, trickle-down theory assumes that an increase in income of the upper class would stimulate the investment thus resulting in

a fall of unemployment. This situation would benefit the general welfare. Then, the trickle-up effect is related with the following expression,

$$\frac{\partial I}{\partial t} = \frac{\frac{A}{N}L}{N(1-r)}, \tag{3.3}$$

and the trickle-down theory with,

$$\frac{\partial I}{\partial t} = \frac{\frac{A}{N}(1-L)}{Nr}, \tag{3.4}$$

where I is the income of a family at rank r, $\frac{A}{N}L$ is the aggregate income of poorer citizens, $N(1-r)$ is the number of wealthy individuals, $\frac{A}{N}(1-L)$ is the income of the wealthiest citizens and Nr is the number of individuals of lower rank. Note that following Equation (3.3), the increase in income of any individual of the society is directly associated with an improvement of the economic situation of low rank citizens. Conversely, Equation (3.4) shows an increment of the upper class income would lead an economic progress of the society as a whole. Finally, the curves (3.1) and (3.2) are the solutions of equations (3.3) and (3.4).

The second motivation is based on Lemma 2.1. If $L_1(p) = p^{1/a}$ is the potential Lorenz curve and $L_2(p) = 1 - (1-p)^a$ with $0 < a \leq 1$ is the Lorenz curve corresponding to the classical Pareto distribution, the family (3.1) is $L_1(L_2(p))$, which is a genuine Lorenz curve according to Lemma 2.1. The second curve (3.2) is of the form $\tilde{L}_1(\tilde{L}_2(p))$, where $\tilde{L}_1(p) = 1 - (1-p)^{1/a}$ and $\tilde{L}_2(p) = p^a$, where $a \geq 1$.

4 The underlying CDF

The underlying cumulative distribution functions (CDF) are obtained in a closed form in the following theorem.

Theorem 4.1 *The underlying CDF associated to the Lorenz curves (3.1) and (3.2) are given by,*

$$F_1(x; a, \mu) = 1 - \frac{1}{\left[1 + (x/\mu)^{a/(1-a)}\right]^{1/a}}, \quad 0 \leq x < \infty, \tag{4.1}$$

and $F_1(x; a, \mu) = 0$ if $x < 0$ and

$$F_2(x; a, \mu) = \frac{1}{\left[1 + (\mu/x)^{a/(a-1)}\right]^{1/a}}, \quad 0 \le x < \infty, \tag{4.2}$$

and $F_2(x; a, \mu) = 0$ if $x < 0$, respectively, where μ represent the mathematical expectation of the population.

Proof: For the LC (3.1) we have,

$$F_1^{-1}(x) = \mu(1-x)^{a-1}[1 - (1-x)^a]^{(1-a)/a},$$

with support $(0, \infty)$. The change of variable $z = (1-x)^a$ leads to $F_1^{-1}(z) = \mu\left(\frac{1-z}{z}\right)^{(1-a)/a}$. Finally, solving previous equation for z and then for x we obtain (4.1). Using a similar reasoning with (3.2), we obtain (4.2). ∎

Both distributions correspond to well known distributions used in the income and wealth literature. The family (4.1) is a Singh-Maddala distribution (Singh and Maddala, 1976) with shape parameters $\frac{a}{1-a}$ and $\frac{1}{a}$ and scale parameter μ. The raw moments of (4.1) are,

$$E(X_1^k) = \frac{\mu^k \Gamma\left(1 + \frac{k(1-a)}{a}\right) \Gamma\left(\frac{1-k(1-a)}{a}\right)}{\Gamma\left(\frac{1}{a}\right)},$$

if $a > 1 - 1/k$, where $\Gamma(x) = \int\limits_0^\infty t^{x-1} e^{-t} dt$ denotes the gamma function. Note that $E(X_1) = \mu$. The mode is at,

$$\mu\left(\frac{2a-1}{2-a}\right)^{(1-a)/a},$$

if $a > 1/2$.

The family (4.2) is a Dagum (1977) distribution with shape parameters $\frac{a}{a-1}$ and $\frac{1}{a}$ and scale parameter μ. In this case, the raw moments are

$$E(X_2^k) = \frac{\mu^k \Gamma\left(\frac{1+k(a-1)}{a}\right) \Gamma\left(1 - \frac{k(a-1)}{a}\right)}{\Gamma\left(\frac{1}{a}\right)},$$

if $a > k/(k-1)$. The mode is at,

$$\mu\left(\frac{2-a}{2a-1}\right)^{(a-1)/a},$$

if $a < 2$.

5 Lorenz ordering, Inequality and Polarization Measures

5.1 Lorenz ordering

The study of Lorenz ordering is a crucial aspect in the analysis of income and wealth distributions. Let $\mathcal{L}$ be the class of all non-negative random variables with positive finite expectation. The Lorenz partial order $\leq_L$ on the class $\mathcal{L}$ is defined by,

$$X \leq_L Y \iff L_X(p) \geq L_Y(p), \quad \forall p \in [0,1].$$

If $X \leq_L Y$, then X exhibits less inequality than Y in the Lorenz sense. We shall show that families (3.1) and (3.2) are ordered with respect a parameters.

Lemma 5.1 *If $L_1(p;a)$ is defined in (3.1) and if $a_1 \leq a_2$, then $L_1(p;a_1) \leq L_1(p;a_2)$, for $0 \leq p \leq 1$. On the other hand, if $a_1 \leq a_2$ then $L_2(p;a_1) \geq L_2(p;a_2)$, for $0 \leq p \leq 1$.*

Proof: If $0 < a \leq 1$, $L_1(p;a)$ is a differentiable function with respect to a, in consequence:

$$\frac{\partial L_1(p;a)}{\partial a} = -L_1(p;a) \left\{ \frac{\log[1-(1-p)^a]}{a^2} + \frac{(1-p)^a \log(1-p)}{a[1-(1-p)^a]} \right\} > 0,$$

for all $p \in (0,1)$. For the $L_2(p;a)$ curve,

$$\frac{\partial L_2(p;a)}{\partial a} = (1-p^a)^{1/a} \left\{ \frac{p^a \log p}{a(1-p^a)} + \frac{\log(1-p^a)}{a^2} \right\} < 0,$$

for all $p \in (0,1)$. ∎

5.2 Gini index

Lemma 5.2 *The Gini indices of the curves $L_1(p,a)$ and $L_2(p,a)$ are given by,*

$$G_1(a) = 1 - \frac{\Gamma(1/a)^2}{a\Gamma(2/a)}, \quad 0 < a \leq 1, \tag{5.1}$$

and

$$G_2(a) = \frac{\Gamma(1/a)^2}{a\Gamma(2/a)} - 1, \quad a \geq 1, \tag{5.2}$$

respectively.

Proof: The Gini index is given by $G = 2 \int_0^1 [p - L(p)]dp = 1 - 2 \int_0^1 L(p)dp$, and substituting (3.1) and (3.2) in previous equation, we obtain easily (5.1) and (5.2), respectively. ∎

Some configurations of the Gini index deserve our attention. If $a = \frac{1}{n}$, with $n = 1, 2, \ldots$ the Gini index (5.1) can be written in the simple form,

$$G_1(n) = 1 - \frac{(n!)^2}{n(2n-1)!},$$

which is of the form $\frac{m}{m+1}$, where m is an integer. For example, for $n = 1, 2, \ldots, 10$ the list of the Gini indices are,

$$0, \quad \frac{2}{3}, \quad \frac{9}{10}, \quad \frac{34}{35}, \quad \frac{125}{126}, \quad \frac{461}{462}, \quad \frac{1715}{1716}, \quad \frac{6434}{6435}, \quad \frac{24309}{24310}, \quad \frac{92377}{92378}.$$

5.3 Donaldson-Weymark-Kakwani index

An important generalization of the Gini index was proposed by Donaldson and Weymark (1980) and Kakwani (1980) and studied by Yitzhaki (1983). These authors proposed the generalized Gini index defined as,

$$G_X(\nu) = 1 - \nu(\nu+1) \int_0^1 (1-p)^{\nu-1} L_X(p)dp, \tag{5.3}$$

where $\nu > 1$ and $L_X(\cdot)$ is the Lorenz curve. If we set $\nu = 1$ in (5.3) we obtain the Gini index. When ν increases, higher weights are attached to small incomes. The limit case when ν goes to infinity depends on the lowest income, expressing the judgement introduced by Rawls, that social welfare depends only on the poorest society member.

Lemma 5.3 *The Donaldson-Weymark-Kakwani indices of the curves $L_1(p, a)$ and $L_2(p, a)$ are given by,*

$$G_{L_1}(\nu; a) = 1 - \nu(\nu+1) \sum_{k=0}^{\infty} \frac{\Gamma(k - 1/a)}{\Gamma(k+1)\Gamma(-1/a)} \frac{1}{\nu + ak}, \tag{5.4}$$

8

and

$$G_{L_2}(\nu; a) = 1 - \nu(\nu+1) \left(\frac{1}{\nu} - \sum_{k=0}^{\infty} \frac{\Gamma(k - 1/a)}{\Gamma(k+1)\Gamma(-1/a)} B(ak+1, \nu) \right) \quad (5.5)$$

respectively, where $B(x,y) = \int_0^1 t^{x-1}(1-t)^{y-1}dt$ *denotes the beta function.*

Proof: Using general expression (5.3) and the series expansion of the binomial expression. ∎

In the case of the curve (3.1), if ν is an integer, expression (5.4) can be written in a simple form. It can be proved (Muliere and Scarsini, 1989) that Equation (5.3) is equivalent to,

$$G_X(\nu) = 1 - \frac{E(X_{1:\nu})}{E(X)}, \quad (5.6)$$

where where $X_{1:\nu}$ represent the minimum random variable in a random sample of size ν. In our case, the distribution of the minimum of a Singh-Maddala distribution is again a Singh-Maddala distribution with parameters $\frac{a}{1-a}$, $\frac{\nu}{a}$ and μ. Using (5.6), we obtain the expression,

$$G_{L_1}(\nu; a) = 1 - \frac{\Gamma\left(\frac{1}{a}\right)\Gamma\left(\frac{\nu-1+a}{a}\right)}{\Gamma\left(\frac{\nu}{a}\right)}.$$

Other simple expressions can also been obtained in the case of L_2. For example, if $\nu = 2$, equation (5.5) can be written of the form,

$$G_{L_2}(2; a) = 1 - 3\left[1 - \Gamma\left(1 + \frac{1}{a}\right)\left(\frac{2^{(a-2)/a}\sqrt{\pi}}{\Gamma\left(\frac{2+a}{2a}\right)} - \frac{\Gamma\left(\frac{2+a}{a}\right)}{\Gamma\left(\frac{3+a}{a}\right)}\right)\right],$$

and if $\nu = 3$, equation (5.5) becomes in,

$$G_{L_2}(3; a) = 1 - 4\left[1 - 3\Gamma\left(1 + \frac{1}{a}\right)\left(\frac{\Gamma\left(\frac{1+a}{a}\right)}{\Gamma\left(\frac{2+a}{a}\right)} - \frac{\Gamma\left(\frac{2+a}{a}\right)}{\Gamma\left(\frac{3+a}{a}\right)} + \frac{\Gamma\left(\frac{3+a}{a}\right)}{3\Gamma\left(\frac{4+a}{a}\right)}\right)\right].$$

5.4 Pietra index

The Pietra index is defined as the maximal vertical deviation between the LC and the egalitarian line

$$P_L = \max_{0 \leq p \leq 1} \{p - L_X(p)\}.$$

If we assume that F_X is strictly increasing on its support, the function $p - L_X(p)$ will be differentiable everywhere on $(0, 1)$ and its maximum will be reached when $1 - F_X^{-1}(x)/\mu_X$ is zero, that is, when $x = F_X(\mu_X)$. The value of $p - L_X(p)$ in this point is given by

$$P_L = F_X(\mu_X) - L_X(F_X(\mu_X)). \tag{5.7}$$

This lemma provides a simple expression for the Pietra's indices of the curves L_1 and L_2.

Lemma 5.4 *The Pietra indices of the curves $L_1(p, a)$ and $L_2(p, a)$ are given by,*

$$P_1(a) = 1 - \frac{1}{2^{1/a-1}}, \quad 0 < a \leq 1, \tag{5.8}$$

and

$$P_2(a) = \frac{1}{2^{1/a-1}} - 1, \quad a \geq 1, \tag{5.9}$$

respectively.

Proof: The proof is direct using formula (5.7) and the expressions (4.1) and (4.2). ∎

5.5 Polarization index

The polarization measurement has been recently proposed as an important variable to characterize income and wealth distributions. It is well known, that polarization is widely accepted as a different concept from inequality. A polarization measure concentrates the income distribution on several polar modes. On the other hand, inequality relates to the overall dispersion of the distribution. A more bipolarized income and wealth distribution is one that is more spread out from the middle, thus implying that the middle class is made up of fewer individuals (Wolfson, 1994).

The Wolfson's index of bipolarization was originally proposed for a population divided in two groups by the median value. This measure is given by,

$$W_X = \frac{2(1/2 - L_X(1/2)) - G_X}{m_X/2\mu_X}, \qquad (5.10)$$

where G_X, m_X and μ_X represent the Gini index, median and mean associated to the Lorenz curve L_X, respectively.

We have the following lemma.

Lemma 5.5 *The Wolfson polarization measures corresponding to the Lamé Lorenz curves (3.1) and (3.2) are given by,*

$$W_{X_1} = 2\frac{\frac{\Gamma(1/a)^2}{2\Gamma(2/a)} - 2(1 - 1/2^a)^{1/a}}{(2^a - 1)^{(1-a)/a}}, \qquad (5.11)$$

and

$$W_{X_2} = 2\frac{2(1 - 1/2^a)^{1/a} - \frac{\Gamma(1/a)^2}{2\Gamma(2/a)}}{(2^a - 1)^{(1-a)/a}}, \qquad (5.12)$$

respectively.

Proof: Previous expressions are obtained from (5.10), using (3.1) and (3.2) together with (4.1) and (4.2) for obtaining the median. ∎

6 Extensions

Recent research about the Lorenz curve (Basmann et al., 1990; Ryu and Slottje, 1996; Sarabia et al., 2005 and Sarabia, 2008a) has shown that some families of LCs approximate only some segments of the income distribution but not others. This fact justifies the consideration of more complex models for the LC beginning with an initial LC.

The following two direct extensions of (3.1) and (3.2) can be considered,

$$L_3(p; a, b) = [1 - (1 - p)^a]^b, \ \ 0 \le p \le 1, \qquad (6.1)$$

and

$$L_4(p; a, b) = 1 - (1 - p^a)^b, \ \ 0 \le p \le 1. \qquad (6.2)$$

Curve (6.1) is genuine if $0 < a \le 1 \le b$ and (6.2) is genuine if $a \ge 1$ and $0 < b \le 1$. Curve (6.1) was proposed by Rasche et al. (1980) and (6.2) was

considered (without to be studied) by Arnold (1987). The Gini index of the Lorenz curve (6.2) is,

$$G_{L_4}(a,b) = \frac{2}{a} \frac{\Gamma\left(\frac{1}{a}\right)\Gamma(b+1)}{\Gamma\left(\frac{1}{a}+b+1\right)} - 1,$$

which will be used in next Section. Other kind extensions can also be considered using the methodologies proposed by Sarabia et al. (1999) and Sarabia et al. (2005).

7 Estimation

In this section we consider two different estimation methods for two different data configuration.

7.1 Estimation from data of the Lorenz curve

Let us suppose that we wish to estimate the parameter a from the curves L_1 and L_2. Let $X_1, \ldots, X_n$ be a sample of size n of income data. The observations consist of n pairs of points $(p_1, q_1), \ldots, (p_n, q_n)$, where $p_i = i/n$, $q_i = s_i/s_n$, and $s_i = x_{1:n} + \cdots + x_{i:n}$ for $i = 1, 2, \ldots, n$, being $x_{i:n}$, $i = 1, 2, \ldots, n$ the i-th order statistic.

The simplest way of estimating a parameter is minimizing the quantity:

$$u(a) = \sum_{i=1}^{n} \left[q_i - L_k(p_i; a)\right]^2, \quad k = 1, 2, \tag{7.1}$$

or alternatively,

$$u_1(a) = \sum_{i=1}^{n} \left[1 - q_i^a - (1-p_i)^a\right]^2, \tag{7.2}$$

for L_1 and

$$u_2(a) = \sum_{i=1}^{n} \left[(1-q_i)^a - 1 + p_i^a\right]^2, \tag{7.3}$$

for L_2, using a non linear optimization standard algorithm, which is available in any of the existing statistical and econometric software packages, including SAS, SPSS, SHAZAM, EViews and Mathematica. We can take $a_0 = 1$ as initial value. An alternative robust method of estimation is given by Castillo, Hadi and Sarabia (1998).

7.2 Estimation with limited information

The practical use of the Lorenz curve requires information on the income and wealth distribution or at least, the income shares for a number of income classes. Unfortunately, sometimes this information is not available. However, in some economic database the available comparable cross-country information is limited to per capita income and the Gini index (see for example Chotikapanich et al, 1997).

Then, assuming that the only available information for the estimation of (3.1) and (3.2) (or (4.1) and (4.2)) is the mean and the Gini index. A plausible estimation method which gives place to consistent estimates, consists of solving the system:

$$\mu = \bar{X},$$
$$G_k(a) = g, \quad k = 1, 2,$$

where $\bar{X}$ and g represent the mean and the Gini sample values, respectively and $G_k(a)$, $k = 1, 2$ the theoretical Gini indices (5.1) and (5.2), respectively. Therefore, the estimate values of μ and a are given by,

$$\hat{\mu} = \bar{X},$$
$$\hat{a} = G_k^{-1}(g), \quad k = 1, 2. \tag{7.4}$$

Note that the right hand side of (7.4) is a monotonic function of parameter g, and consequently has only one solution.

8 Empirical Applications

8.1 Estimation when some point of the Lorenz curve are available

Results of the estimation of the Lorenz curves studied and the corresponding Gini index are presented in this section. The source of data used for this purpose is taken from Shorrocks (1983). The data correspond to cumulated income shares for 19 countries derived from Jain (1975). This data set is relevant since the sample considered is characterized by heterogeneity among countries in terms of income inequality, thus allowing us to draw conclusions regarding the adjustment of the Lorenz curves for different levels of inequality.

Tables 1 to 4 present the estimation results for the two Lamé Lorenz curves (3.1) and (3.2) and for the two extensions (6.1) and (6.2), respectively. The estimators have been obtained by non-linear least squares, according to the method established in the previous section, minimizing expression (7.1) or (7.2) and (7.3). Using a consistent estimate of the covariance matrix of the coefficients (Amemiya, 1985) we computed the corresponding standard errors. These tables also give several error measures. The mean square error (MSE), which is given by,

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^{n} \left[q_i - L(p_i; \hat{\theta}) \right]^2,$$

the mean absolute error (MAE),

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^{n} \left| q_i - L(p_i; \hat{\theta}) \right|,$$

and the maximum absolute error (MAX),

$$\text{MAX} = \max_{i=1,\dots,n} \left| q_i - L(p_i; \hat{\theta}) \right|,$$

where $L(p; \hat{\theta})$ represents the Lorenz curve estimated. It can be concluded that the empirical results reported in this study indicate that the four functions, though simple and easy to estimate, are very satisfactory in fitting data. We also present the estimates of the Gini indices associated with the four LCs under study, which are presented in the last column of each table.

From Tables 1 and 2 we observe that almost 50 percent of the countries considered are better fitted by the curve L_1, whereas the curve L_2 outperforms the adjustment in the rest of the countries. As consequence, we can not conclude the superiority of one curve over the other, when the previous three error criteria are considered.

If we consider the two extensions L_3 and L_4, the superiority of the curve L_4 is concluded, since it adjust the income distribution more adequately in 13 out of 19 countries included in the sample (Tables 3 and 4). It is important to note that all of these countries are characterized by low levels of inequality (Gini values lower than 0.5). On the contrary, the L_3 LC fits better income distributions of countries with high inequality (Gini values greater than 0.5).

Regarding the comparison of the general models L_3 and L_4 (Equations (3.1) and (3.2)) with its special cases L_1 and L_2 (Equations (6.1) and (6.2)), some comments are included. As might be expected, the extensions considered outperform single-parameter curves since the three error measures are substantially lower for the curves L_4 and L_3. However, Gini index estimates based on Lamé curves and its extensions are practically identical, differing by no more than 0.035 and, in the majority of cases the pairwise difference is lower than 0.01. In line with other studies that compare the performance of more complex functional forms with one-parameter curves (Ogwang and Rao, 1996), we observe that the greater differences correspond to the countries that present higher levels of inequality. Thus, following the parsimony principle, the use of more complicated expressions would not be relevant for the adjustment of the Lamé Lorenz curves and the posterior estimation of the corresponding Gini index, at least for the data set considered.

8.2   Estimation with limited information

In this section we focus on international inequality at global level, using the methodology proposed in Section 7. We use the information provided by Bourguignon and Morrison (2002), which includes the estimates of several world inequality measures. Using the empirical Gini index and equation (7.4) with L_1 and L_2, we have obtained the single-parameter estimates of the inequality indices a. The standard errors were computed using parametric bootstrap with $B = 999$ bootstrap replications. Using these estimators, we have obtain the Pietra and Polarization indices given in equations (5.8)-(5.9) and (5.11)-(5.12), respectively. All these estimators are recorded in Tables 5 and 6.

The results point out that inequality has increased in the last two centuries, conclusion that is congruent with the evolution of the Gini index. Moreover, the polarization has intensified over the study period thus indicating that the differences among wealthiest and poorer countries have increased.

9   Relationships with other curves

Bonferroni (1930) considered a curve for studying inequality, which is more suitable for the analysis of low income groups. The curve, in terms of the

quantile function is given by,

$$B_X(p) = \frac{1}{p\mu_X} \int\limits_0^p F_X^{-1}(u)du = \frac{L_X(p)}{p}, \ \ 0 \le p \le 1. \tag{9.1}$$

In the context of informetric data, Sarabia (2008b) has proposed another curve namely the Leimkuhler curve (see also Balakrishnan et al (2010) and Sarabia and Sarabia (2008)). This curve is defined as

$$K_X(p) = \frac{1}{\mu_X} \int\limits_{1-p}^1 F_X^{-1}(u)du = 1 - L_X(1-p), \ \ 0 \le p \le 1. \tag{9.2}$$

In our case, the Bonferroni and Leimkuhler curves associated to (3.1) and (3.2) can be obtained directly using expressions (9.1) and (9.2).

Acknowledgements

References

Aggarwal, V. (1984). On Optimum Aggregation of Income Distribution Data. *Sankhya B* 46:343-355.

Aggarwal, V., Singh, R. (1984). On Optimum Stratification with Proportional Allocation for a Class of Pareto Distributions. *Communications in Statistics, Theory and Methods* 13:3017-3116.

Amemiya, T. (1985). Advanced Econometrics, Oxford: Basic Blackwell.

Arnold, B.C. (1986). A Class of Hyperbolic Lorenz Curves. *Sankhya B* 48:427-436.

Arnold, B.C.(1987). *Majorization and the Lorenz Curve: a Brief Introduction*: Vol. 43. Lecture Notes in Statistics, Berlin: Springer-Verlag.

Arnold, B.C., Robertson, C.A., Brockett, P.L., and Shu, B.Y. (1987). Generating Ordered Families of Lorenz Curves by Strongly Unimodal Distributions. *Journal of Business and Economic Statistics* 5:305-308.

Balakrishnan, N., Sarabia, J.M., Kolev, N. (2010). A Simple Relation Between the Leimkuhler Curve and the Mean Residual Life. *Journal of Informetrics* 4: 602-607.

Basmann, R.L., Hayes, K.L., Slottje, D.J., J.D. Johnson (1990). A General Functional Form for Approximating the Lorenz Curve. *Journal of Econometrics* 92(4):727-744.

Bonferroni, C.E. (1930). *Elementi di Statistica Generale.* Seeber, Firenze.

Bourguignon, F., Morrisson, C. (2002). Inequality among world citizens: 1820-1992. *The American Economic Review* 92: 727-744.

Castillo, E., Hadi, A.S., Sarabia, J.M. (1998). A Method for Estimating Lorenz Curves. *Communications in Statistics, Theory and Methods* 27:2037-2063.

Chotikapanich, D. (1993). A Comparison of Alternative Functional Forms for the Lorenz Curve. *Economics Letters* 41:129-138.

Chotikapanich, D., Valenzuela, R., Rao, D.S.P. (1997). Global and regional Inequality in the Distribution of Income: Estimation with Limited and Incomplete Data. *Empirical Economics* 22:533-546.

Dagum, C. (1977). A new model a personal income distribution: Specification and estimation. *Economie Appliquée* 30:413-437.

Donaldson, D., and Weymark, J.A. (1980). A single-parameter generalization of the Gini index of inequality. *Journal of Economic Theory* 22: 67-86.

Gastwirth, J.L. (1971). A General Definition of the Lorenz Curve. *Econometrica* 39:1037-1039.

Gastwirth, J.L. (1972). The Estimation of the Lorenz Curve and Gini Index. *The Review of Economics and Statistics* 54:306-316.

Helene, O. (2010). Fitting Lorenz curves. *Economics Letters* 108:153-155.

Henle, J.M., Horton, N.J., Jakus, S.J. (2008). Modelling Inequality with a Single Parameter. In: *Modeling Income Distributions and Lorenz Curves*: Vol. 5, Economic Studies in Inequality, Social Exclusion and Well-Being, D. Chotikapanich Edit., Springer Verlag, New York.

Holm, J. (1993). Maximum Entropy Lorenz Curves. *Journal of Econometrics* 44:377-389.

Jain, S. (1975). *Size Distribution of Income: A Compilation of Data*. Washington: World Bank.

Johnson, N.L., Kotz, S. Balakrishnan, N. (1994). *Continuous Univariate Distributions*, Vol. 1. Second Edition. New York: John Wiley.

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

Kakwani, N.C., Podder, N. (1976). Efficient Estimation of the Lorenz Curve and Associated Inequality Measures from Grouped Observations. *Econometrica* 44:137-148.

Muliere,P., Scarsini, M. (1989). A Note on Stochastic Dominance and Inequality Measures. *Journal of Economic Theory* 49:314-323.

Ogwang, T., Rao, U.L.G. (1996). A new functional form for approximating the Lorenz curve. *Economics Letters* 52:21-29.

Ogwang, T., Rao, U.L.G. (2000). Hybrid models of the Lorenz curve. *Economics Letters* 69:39-44.

Ortega, P., Martín, A., Fernández, A., Ladoux, M., A. García (1991). A New Functional Form for Estimating Lorenz Curves. *Review of Income and Wealth* 37:447-452.

Pakes, A. G. (1986). On Income Distributions and Their Lorenz Curves. *Technical Report, Department of Mathematics, University of Western Australia*, Nedlands, W.A.

Rasche, R.H., Gaffney, J., Koo, A.Y.C., N. Obst. (1980). Functional Forms for Estimating the Lorenz Curve. *Econometrica* 48:1061-1062.

Rhode, N. (2009). An alternative functional form for estimating the Lorenz curve. *Economics Letters* 105: 61-63.

Ryu, H. and Slottje, D.J. (1996). Two Flexible Functional Forms for Approximating the Lorenz Curve. *Journal of Econometrics* 72:251-274.

Sarabia, J.M. (1997). A Hierarchy of Lorenz Curves based on the Generalized Tukey's Lambda Distribution. *Econometrics Reviews* 16:305-320.

Sarabia, J.M. (2008a). Parametric Lorenz Curves: Models and Applications. In: *Modeling Income Distributions and Lorenz Curves*, D. Chotikapanich Edit., Economic Studies in Inequality, Social Exclusion and Well-Being, Vol. 5:67-190. New York: Springer Verlag.

Sarabia, J.M. (2008b). A general definition of Leimkuhler curve. *Journal of Informetrics* 2:56-163.

Sarabia, J.M., Castillo, E., Pascual, M., Sarabia, M. (2005). Mixture Lorenz curves. *Economics Letters* 89:89-94.

Sarabia, J.M., Castillo, E., Slottje, D.J. (1999). An Ordered Family of Lorenz Curves. *Journal of Econometrics* 91:43-60.

Sarabia, J.M., Castillo, E., Slottje, D.J. (2001). An Exponential Family of Lorenz Curves. *Southern Economic Journal* 67: 748-756

Sarabia, J.M., Pascual, M. (2002). A Class of Lorenz Curves Based on Linear Exponential Loss Functions. *Communications in Statistics, Theory and Methods* 31:925-942.

Sarabia, J.M., Prieto, F., Sarabia, M. (2010). Revisiting a functional form for the Lorenz curve. *Economics Letters*, 107:249-252.

Sarabia, J.M., Sarabia, M. (2008). Explicit expressions for the Leimkuhler curve in parametric families. *Information Processing and Management* 44: 1808-1818.

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

Singh, S.K., Maddala, G.S. (1976). A function for the size distribution of incomes. *Econometrica* 44:963-970.

Slottje, D. (2010). Modeling income distributions and Lorenz curves. *Journal of Economic Inequality* 8:525-526.

Villaseñor, J.A., Arnold B.C. (1989). Elliptical Lorenz Curves. *Journal of Econometrics* 40:327-338.

Wolfson, M.C. (1994). When inequalities diverge. *American Economic Review*: 84:353-358.

Yitzhaki, S. (1983). On a Extension of the Gini Inequality Index. *International Economic Review* 24:617-628.

Table 1: Fitted Lorenz Curves, error measures and Gini index for the Lorenz curve (3.1). Standard errors in parentheses.

Country	a	MSE$\times 10^2$	MAE	MAX	Gini index
Brazil	0.5051	0.0824	0.0254	0.0592	0.6587
	(0.0091)				
Columbia	0.5646	0.0230	0.0125	0.0343	0.5658
	(0.0051)				
Denmark	0.7013	0.0170	0.0111	0.0288	0.3609
	(0.0053)				
Finland	0.6314	0.0142	0.0109	0.0191	0.4634
	(0.0044)				
India	0.6246	0.1010	0.0262	0.0701	0.4736
	(0.0117)				
Indonesia	0.6264	0.2341	0.0408	0.1040	0.4709
	(0.0178)				
Japan	0.7610	0.0008	0.0023	0.0050	0.2785
	(0.0012)				
Kenya	0.5072	0.1537	0.0355	0.0765	0.6555
	(0.0124)				
Malaysia	0.5980	0.0042	0.0052	0.0152	0.5142
	(0.0023)				
Netherlands	0.6426	0.0005	0.0018	0.0044	0.4466
	(0.0008)				
New Zealand	0.6997	0.0144	0.0110	0.0170	0.3632
	(0.0049)				
Norway	0.7067	0.0183	0.0125	0.0186	0.3533
	(0.0056)				
Panama	0.6430	0.0006	0.0020	0.0048	0.4460
	(0.0009)				
Sri Lanka	0.6690	0.0027	0.0041	0.0116	0.4075
	(0.0020)				
Sweden	0.6859	0.0059	0.0068	0.0151	0.3830
	(0.0030)				
Tanzania	0.5652	0.1258	0.0314	0.0680	0.5649
	(0.0120)				
Tunisia	0.6036	0.0102	0.0087	0.0217	0.5056
	(0.0036)				
United Kingdom	0.6996	0.0034	0.0043	0.0140	0.3634
	(0.0023)				
Uruguay	0.6138	0.0274	0.0132	0.0378	0.4900
	(0.0059)				

Table 2: Fitted Lorenz Curves, error measures and Gini index for the Lorenz curve (3.2). Standard errors in parentheses.

Country	a	MSE$\times 10^2$	MAE	MAX	Gini index
Brazil	2.3455	0.0745	0.0245	0.0528	0.6594
	(0.0627)				
Columbia	1.9856	0.0185	0.0155	0.0284	0.5664
	(0.0234)				
Denmark	1.4853	0.0191	0.0116	0.0312	0.3611
	(0.0149)				
Finland	1.7006	0.0161	0.0115	0.0233	0.4637
	(0.0169)				
India	1.7262	0.0932	0.0257	0.0657	0.4743
	(0.0417)				
Indonesia	1.7197	0.2228	0.0403	0.0994	0.4717
	(0.0640)				
Japan	1.4000	0.0011	0.0027	0.0058	0.3125
	(0.0032)				
Kenya	2.3306	0.1442	0.0346	0.0701	0.6562
	(0.0862)				
Malaysia	1.8310	0.0027	0.0044	0.0101	0.5147
	(0.0078)				
Netherlands	1.6617	0.0003	0.0016	0.0032	0.4470
	(0.0022)				
New Zealand	1.4899	0.0155	0.0113	0.0195	0.3635
	(0.0135)				
Norway	1.4715	0.0188	0.0127	0.0176	0.3535
	(0.0146)				
Panama	1.6601	0.0001	0.0008	0.0032	0.4463
	(0.0016)				
Sri Lanka	1.5766	0.0042	0.0047	0.0148	0.4078
	(0.0077)				
Sweden	1.5773	0.0072	0.0073	0.0180	0.4081
	(0.0096)				
Tanzania	1.9826	0.1184	0.0308	0.0622	0.5655
	(0.0589)				
Tunisia	0.8071	0.0151	0.0106	0.0264	0.5059
	(0.0181)				
United Kingdom	1.4904	0.0023	0.0037	0.0113	0.3638
	(0.0052)				
Uruguay	1.7662	0.0331	0.0142	0.0423	0.4903
	(0.0258)				

Table 3: Fitted Lorenz Curves, error measures and Gini index for the Lorenz curve (6.1). Standard errors in parentheses.

Country	a	b	MSE$\times 10^2$	MAE	MAX	Gini index
Brazil	0.3580	1.3803	0.0021	0.0037	0.0090	0.6348
	(0.0078)	(0.0274)				
Columbia	0.4868	1.5033	0.0018	0.0037	0.0086	0.5552
	(0.0079)	(0.0253)				
Denmark	0.7806	1.6136	0.0017	0.0031	0.0100	0.3670
	(0.0098)	(0.0239)				
Finland	0.7015	1.7930	0.0004	0.0016	0.0043	0.4702
	(0.0046)	(0.0145)				
India	0.4647	1.1806	0.0901	0.0080	0.0187	0.4551
	(0.0167)	(0.0395)				
Indonesia	0.3912	1.0088	0.0167	0.0110	0.0250	0.4424
	(0.0207)	(0.0447)				
Japan	0.7235	1.3330	0.0004	0.0017	0.0029	0.3114
	(0.0045)	(0.0090)				
Kenya	0.3114	1.2095	0.0020	0.0037	0.0083	0.6217
	(0.0070)	(0.0226)				
Malaysia	0.5643	1.5661	0.0006	0.0020	0.0046	0.5100
	(0.0047)	(0.0142)				
Netherlands	0.6463	1.5665	0.0005	0.0016	0.0041	0.4470
	(0.0045)	(0.0124)				
New Zealand	0.7751	1.6075	0.0005	0.0014	0.0057	0.3688
	(0.0051)	(0.0126)				
Norway	0.7893	1.6060	0.0018	0.0035	0.0092	0.3590
	(0.0104)	(0.0250)				
Panama	0.6401	1.5469	0.0006	0.0020	0.0047	0.4456
	(0.0051)	(0.0139				
Sri Lanka	0.6965	1.5651	0.0007	0.0023	0.0043	0.4100
	(0.0057)	(0.0148)				
Sweden	0.7325	1.5713	0.0003	0.0015	0.0034	0.3867
	(0.0040)	(0.0099)				
Tanzania	0.3859	1.1960	0.0012	0.0029	0.0063	0.5378
	(0.0056)	(0.0154)				
Tunisia	0.6245	1.7243	0.0089	0.0082	0.0155	0.5081
	(0.0197)	(0.0640)				
United Kingdom	0.6725	1.3691	0.0014	0.0032	0.0074	0.3613
	(0.0081)	(0.0177)				
Uruguay	0.7056	1.9287	0.0040	0.0053	0.0118	0.5000
	(0.0142)	(0.0491)				

Table 4: Fitted Lorenz Curves, error measures and Gini index for the Lorenz curve (6.2). Standard errors in parentheses.

Country	a	b	MSE$\times 10^2$	MAE	MAX	Gini index
Brazil	1.5694	0.3193	0.0012	0.0027	0.0063	0.6362
	(0.0296)	(0.0042)				
Columbia	1.6846	0.4430	0.0006	0.0019	0.0054	0.5567
	(0.0183)	(0.0037)				
Denmark	1.6858	0.7547	0.0024	0.0038	0.0116	0.3675
	(0.0291)	(0.0122)				
Finland	1.9410	0.6596	0.0003	0.0016	0.0026	0.4711
	(0.0129)	(0.0039)				
India	1.2429	0.4479	0.0073	0.0071	0.0176	0.4560
	(0.0455)	(0.0123)				
Indonesia	1.0223	0.3927	0.0164	0.0108	0.0260	0.4432
	(0.0596)	(0.0159)				
Japan	1.3774	0.7041	0.0008	0.0025	0.0040	0.3117
	(0.0136)	(0.0061)				
Kenya	1.3159	0.2882	0.0015	0.0034	0.0065	0.6226
	(0.0297)	(0.0043)				
Malaysia	1.7280	0.5210	0.0001	0.0008	0.0019	0.5113
	(0.0075)	(0.0019)				
Netherlands	1.6863	0.6093	0.0001	0.0008	0.0026	0.4479
	(0.0068)	(0.0021)				
New Zealand	1.6801	0.7482	0.0003	0.0013	0.0047	0.3694
	(0.0106)	(0.0044)				
Norway	1.6723	0.7637	0.0001	0.0027	0.0083	0.3595
	(0.0202)	(0.0087)				
Panama	1.6644	0.6037	0.0001	0.0008	0.0031	0.4465
	(0.0056)	(0.0023)				
Sri Lanka	1.6622	0.6641	0.0016	0.0036	0.0066	0.4106
	(0.0242)	(0.0086)				
Sweden	1.6551	0.7030	0.0076	0.0023	0.0054	0.3874
	(0.0163	(0.0063)				
Tanzania	1.2745	0.3647	0.0010	0.0028	0.0045	0.5386
	(0.0198)	(0.0040)				
Tunisia	1.9037	0.5776	0.0131	0.0100	0.0187	0.5091
	(0.0873)	(0.0226)				
United Kingdom	1.4328	0.6483	0.0005	0.0022	0.0053	0.3691
	(0.0138)	(0.0054)				
Uruguay	2.1110	0.6599	0.0066	0.0070	0.0146	0.5010
	(0.0657)	(0.0185)				

24

Table 5: Estimated inequality and polarization measures with limited information using the curve (3.1). Standard errors based on parametric bootstrap in parentheses.

Year	Gini index	a	Pietra index	Wolfson index
1820	0.5000	0.6073	0.3613	0.4734
		(0.0425)		
1850	0.5320	0.5864	0.3867	0.5198
		(0.0426)		
1870	0.5600	0.5683	0.4093	0.5642
		(0.0429)		
1890	0.5880	0.5504	0.4324	0.6128
		(0.0436)		
1910	0.6100	0.5363	0.4508	0.6546
		(0.0475)		
1929	0.6160	0.5325	0.4559	0.6666
		(0.0468)		
1950	0.6400	0.5171	0.4765	0.7176
		(0.0456)		
1960	0.6350	0.5203	0.4722	0.7066
		(0.0469)		
1970	0.6500	0.5107	0.4853	0.7405
		(0.0444)		
1992	0.6570	0.5062	0.4914	0.7572
		(0.0491)		

Table 6: Estimated inequality and polarization measures with limited information using the curve (3.2). Standard errors based on parametric bootstrap in parentheses.

Year	Gini index	a	Pietra index	Wolfson index
1820	0.5000	1.7915	0.3583	0.4569
		(0.2258)		
1850	0.5320	1.8800	0.3833	0.5003
		(0.2131)		
1870	0.5600	1.9651	0.4055	0.5416
		(0.3150)		
1890	0.5880	2.0584	0.4282	0.5870
		(0.3151)		
1910	0.6100	2.1386	0.4463	0.6259
		(0.3226)		
1929	0.6160	2.1616	0.4513	0.6371
		(0.3605)		
1950	0.6400	2.2593	0.4716	0.6848
		(0.4649)		
1960	0.6350	2.2382	0.4663	0.6744
		(0.3667)		
1970	0.6500	2.3029	0.4802	0.7062
		(0.3630)		
1992	0.6570	2.3345	0.4862	0.7218
		(0.4138)		