Empir Econ (2016) 51:1215–1233
DOI 10.1007/s00181-015-1031-y

# A family of arctan Lorenz curves

**E. Gómez-Déniz**[1]

Received: 23 May 2014 / Accepted: 20 October 2015 / Published online: 17 November 2015


**Abstract** This paper presents a new family of parametric Lorenz curves based on the arctan function and adding a parameter $-\infty < \alpha < \infty$, $\alpha \neq 0$ to an initial Lorenz curve $L_0(p)$, $0 \leq p \leq 1$. The particular case obtained when $\alpha$ tends to zero is reduced to the initial Lorenz curve $L_0(p)$. The corresponding distribution functions are shown. Some inequality measures are calculated, and a method to compute the Gini index based on the use of the inverse of the Lorenz curve is proposed. Finally, an application to two well-known data sets is presented and a good fit is obtained.

**Keywords** Parametric Lorenz curve · Gini measure · Leimkuhler curve · Pietra measure

**Mathematics Subject Classification** 62P20 · C80 · D30

## 1 Introduction

This paper introduces a parametric family of Lorenz curves obtained by a general method, based on adding a parameter $-\infty < \alpha < \infty$, $\alpha \neq 0$ to an initial Lorenz curve $L_0(p)$, $0 \leq p \leq 1$, using the arctan function. The particular case obtained when $\alpha$ tends to zero is reduced to the initial Lorenz curve $L_0(p)$.

The development of new functional forms of Lorenz curves has been an attractive area of research in recent decades; see, for example, Kakwani (1980), Aggarwal and Singh (1984), Gupta (1984), Ortega et al. (1991), Basmann et al. (1990), Chotikapanich

✉ E. Gómez-Déniz
emilio.gomez-deniz@ulpgc.es

1 Department of Quantitative Methods in Economics, TiDES Institute, University of Las Palmas de Gran Canaria, 35017 Las Palmas de G.C., Spain

(1993), Ogwang and Rao (1996), Sarabia et al. (1999, 2010). For a recent review of Lorenz curves and income distributions, see Chotikapanich (2008). These methods also provide new functional forms of Leimkuhler curves, which are interesting in terms of informetrics and in particular regarding concentration aspects in this field (see Burrell 1992, 2005; Sarabia and Sarabia 2008; Sarabia et al. 2010, among others).

The densities and distribution functions corresponding to the new Lorenz curves and the corresponding Gini and Pietra inequality indices are shown in closed forms for some particular cases. A method based on the use of the inverse of the initial Lorenz curve is given, which facilitates the computation of the Gini index with the family proposed here.

In this study, we use two data sets (1977 and 1990) from the US Current Population Survey (CPS), considered in Ryu and Slottje (1996), and compare the results with those of the initial Lorenz curves examined.

The structure of this paper is as follows. In Sect. 2, we describe the new family of arctan Lorenz curves and the corresponding Leimkuhler curves. Some particular cases obtained by starting with an initial Lorenz curve $L_0(p)$ are shown. In Sect. 3, the Gini and Pietra indices are obtained, together with the population functions for some cases. In Sect. 4, we compare the performance of the proposed Lorenz curves with that of the initial ones by fitting them to the two data sets, and finally, in Sect. 5, our main conclusions are presented.

## 2 The new family of Lorenz curves

This section begins with the definition of the Lorenz curve provided by Gastwirth (1971) in accordance with the original proposal by Pietra (1915). Thus:

**Definition 1** Given a distribution function $F(x)$ with support in the subset of the positive real numbers and with finite expectation $\mu$, we define a Lorenz curve as

$$L_F(p) = \frac{1}{\mu}\int_0^p F^{-1}(x)\mathrm{d}x, \quad 0 \leq p \leq 1, \tag{1}$$

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

A characterization of the Lorenz curve, which is well known in the literature, is given by the following result.

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

$$L(0) = 0, \quad L(1) = 1, \quad L'(0^+) \geq 0 \quad \text{for} \quad p \in (0, 1), \quad L''(p) \geq 0. \tag{2}$$

The main result of this paper is expressed in the following theorem.

**Theorem 2** *Let $L_0(p)$ be a Lorenz curve, $-\infty < \alpha < \infty$, $\alpha \neq 0$, a real parameter and consider the transformation*

$$L_\alpha(p) = 1 - \frac{\arctan\left(\alpha(1 - L_0(p))\right)}{\arctan \alpha}, \quad 0 \leq p \leq 1. \tag{3}$$

*Then, $L_\alpha(p)$ is also a Lorenz curve.*

*Proof* Simple algebra provides that $L_\alpha(0) = 0,\ L_\alpha(1) = 1,$

$$L'_\alpha(p) = \frac{\alpha}{\arctan\alpha}\frac{L'_0(p)}{1+(\alpha(1-L_0(p)))^2} > 0,$$
$$L''_\alpha(p) = \frac{1}{1+(\alpha(1-L_0(p)))^2}\left[\frac{\alpha L''_0(p)}{\arctan\alpha} + 2\alpha^2 L'_0(p)L'_\alpha(p)(1-L_0(p))\right] > 0,$$

and $L_\alpha(p) < p$. Then, if $L_0(p)$ is a genuine Lorenz curve, expression (3) possesses the proper convexity and slope constraints for us to assure that it always lies in the lower triangle of the unit square, and therefore, $L_\alpha(p)$ represents a genuine Lorenz curve. □

Using the well-known result that establishes that

$$\arctan u - \arctan v = \arctan\left(\frac{u-v}{1+uv}\right)$$

(3) can be rewritten in a more compact form as

$$L_\alpha(p) = \frac{1}{\arctan\alpha}\arctan\left(\frac{\alpha L_0(p)}{1+\alpha^2(1-L_0(p))}\right). \tag{4}$$

By taking in (3) or alternatively in (4) the limit when the parameter $\alpha$ tends to zero and applying L'Hospital's rule, it is straightforward to derive that the initial Lorenz curve $L_0(p)$ is obtained as a special case, i.e., $L_\alpha(p) \to L_0(p)$ when $\alpha \to 0$. Thus, the methodology proposed here can be considered as a mechanism for adding a parameter to an initial Lorenz curve and therefore a means of obtaining a more flexible Lorenz curve.

Other ways to write $L_\alpha(p)$ given in (4) can be obtained by using the following representation of the arctan function (see Castellanos 1988):

$$\arctan z = \frac{z}{1+z^2}\,{}_2F_1\left(1, 1; \frac{3}{2}, \frac{z^2}{1+z^2}\right) = \sum_{n=0}^{\infty}\frac{2^{2n}(n!)^2}{(2n+1)!}\frac{z^{2n+1}}{(1+z^2)^{n+1}}. \tag{5}$$

Here ${}_2F_1(a, b; c, z)$ represents the hypergeometric function which has the integral representation

$${}_2F_1(a, b; c, z) = \frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)}\int_0^1 t^{b-1}(1-t)^{c-b-1}(1-tz)^{-a}\,\mathrm{d}t, \tag{6}$$

and where $\Gamma(\cdot)$ is the Euler gamma function.

Approximations to the arctan function can be obtained using second- and third-order polynomials and simple rational functions (see Rajan et al. 2006 for details), and it is thus obtained that $\arctan((1+x)/(1-x)) \approx \pi(x+1)/4$. Applying this to (3) and after some algebra, we have

$$L_\alpha(p) \approx \frac{L_0(p)}{1+\alpha(1-L_0(p))}, \quad \alpha > 0. \tag{7}$$

Observe that the right-hand side in (7) is a genuine Lorenz curve and coincides with expression (27) in Sarabia et al. (2010). Additionally, the Aggarwal and Singh Lorenz curve (see Aggarwal and Singh 1984; Arnold 1986) is obtained from (7) when $L_0(p) = p$. The mechanism proposed here is more general that the one proposed in Sarabia et al. (2010).

Expression (7) can also be obtained by considering the ordered sequence of Lorenz curves given by

$$L_0(p) \geq L_0(p)^2 \geq \cdots \geq L_0(p)^n \geq \dots \tag{8}$$

where $n$ is an integer. It is possible to build a new family of Lorenz curves beginning from (8), but now assuming that the powers $\{1, 2, \dots, n, \dots\}$ are not fixed, and are distributed according to a convenient discrete random variable with probability mass function $P_j = Pr(X = j)$, $j = 1, 2, \dots$. In the particular case that $P_j = 1/(1+\alpha)\,(\alpha/(1+\alpha))^{n-1}$, $\alpha > 0$, i.e., the geometric distribution, the family of Lorenz curves gives (7).

It is known that the Lorenz curve determines the distribution of $X$ up to a scale factor transformation, since $F^{-1}(x) = \mu L'(x)$. Moreover, the relation

$$K_0(p) = 1 - L_0(1-p) \tag{9}$$

determines the relationship between the Lorenz and the Leimkuhler curves (see Sarabia and Sarabia 2008 and Sarabia et al. 2010, among others). This curve plays an important role in informetrics (see, for instance, Burrell 1992, 2005). Therefore, from (3) and (9), we can also define a family of arctan Leimkuhler curves starting from an initial Lorenz curve $L_0(p)$, given by

$$K_\alpha(p) = \frac{\arctan(\alpha(1-L_0(1-p)))}{\arctan\alpha}, \quad -\infty < \alpha < \infty,\ \alpha \neq 0.$$

### 2.1 Lorenz ordering

Lorenz ordering is an important aspect in the analysis of income and wealth distributions. If we define $L$ to be the class of all nonnegative random variables with positive finite expectation, the Lorenz partial order $\leq_L$ on the class $L$ 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. In the next result, we show that family (3) is ordered with respect to parameter $\alpha$.

**Proposition 1** *The Lorenz curve $L_\alpha(p)$ is ordered with respect to $\alpha$, i.e., if $|\alpha_1| \leq |\alpha_2|$, $-\infty < \alpha_1, \alpha_2 < \infty$, $\alpha_1, \alpha_2 \neq 0$, then $L_{|\alpha_1|}(p) \geq L_{|\alpha_2|}(p)$, for $0 \leq p \leq 1$.*

*Proof* After computing the derivative of the logarithm of (3), then the sign of $\mathrm{d}L_\alpha(p)/\mathrm{d}\alpha$ depends on the sign of

$$\begin{aligned}\Phi_\alpha(p) = &-[1 - L_\alpha(p)]\left\{[1 - L_0(p)]\left(1+\alpha^2\right)\arctan\alpha\right.\\ &\left.-\left[1+\alpha^2(1-L_0(p))^2\right]\arctan\left(\alpha\left(1-L_0(p)\right)\right)\right\}.\end{aligned}$$

Now, using the following inequalities

$$\begin{gathered}\left(1+\alpha^2\right)(1-L_0(p)) > \left[1+\alpha^2\left(1-L_0(p)\right)\right][1-L_0(p)],\\ \arctan\alpha > \arctan\left(\alpha\left(1-L_0(p)\right)\right),\end{gathered}$$

it is simple to see that $\Phi_\alpha(p) < 0$.

Hence, the result. □

The following result sustains that the equality is obtained, i.e., $X$ exhibits the same inequality as $Y$, when $\alpha_1 = -\alpha_2$.

**Proposition 2** *It is verified that $L_\alpha(p) = L_{-\alpha}(p)$, for all $-\infty < \alpha < \infty$, $\alpha \neq 0$ and $0 \leq p \leq 1$.*

*Proof* Self-evident. □

### 2.2 New functional forms of Lorenz curves

In order to derive new functional forms of Lorenz curves, we now consider the following initial Lorenz curves: egalitarian, Aggarwal and Singh Lorenz curve and Pareto Lorenz curve.

The arctan egalitarian Lorenz curve is obtained in (4), by replacing the initial Lorenz curve with $L_0(p) = p$. Thus, it is given by

$$L_\alpha(p) = \frac{1}{\arctan\alpha}\arctan\left(\frac{\alpha p}{1+\alpha^2(1-p)}\right), \quad -\infty < \alpha < \infty,\ \alpha \neq 0. \tag{10}$$

The arctan Aggarwal and Singh Lorenz curve is obtained in a similar way, replacing the initial Lorenz curve (see Aggarwal and Singh 1984; Arnold 1986) with $L_0(p) = p/(1+\theta(1-p)), \quad \theta > 0$, and therefore we have

$$L_\alpha(p) = \frac{1}{\arctan\alpha}\arctan\left(\frac{\alpha p}{1+(1-p)(\theta+\alpha^2(1+\theta))}\right), \tag{11}$$

where $\theta > 0,\ -\infty < \alpha < \infty,\ \alpha \neq 0$.

Consider now the Pareto Lorenz curve

$$L_0(p) = 1 - (1-p)^\theta,\ 0 < \theta < 1,$$

from which we obtain the arctan Pareto Lorenz curve

$$L_\alpha(p) = \frac{1}{\arctan\alpha} \arctan\left(\frac{\alpha(1-(1-p)^\theta)}{1+\alpha^2(1-p)^\theta}\right). \tag{12}$$

Finally, by taking as the initial one the Chotikapanich Lorenz curve given by $L_0(p) = (\exp(\theta p) - 1)/(\exp(\theta) - 1),\ \theta > 0$, we obtain the arctan Chotikapanich Lorenz curve

$$L_\alpha(p) = \frac{1}{\arctan\alpha} \arctan\left(\frac{\alpha(\exp(\theta p) - 1)}{\exp(\theta) - 1 + \alpha^2(\exp(\theta) - \exp(\theta p))}\right). \tag{13}$$

Of course, other arctan Lorenz curves can be obtained by replacing $L_0(p)$ in (4) with other initial Lorenz curves, such as the Gupta or generalized Pareto Lorenz curves. We chose the above initial Lorenz curves because, as discussed in the next section, closed-form expressions can be obtained for some inequality measures and population functions.

## 3 Inequality measures and population functions

The corresponding Gini and Pietra indices can be computed straightforwardly when the egalitarian and Aggarwal initial Lorenz curves are chosen as $L_0(p)$.

### 3.1 Gini and Yitzhaki indices

The Gini coefficient (also known as the Lorenz concentration ratio) is a measure (degree of concentration) of the inequality of a variable in a distribution of its elements, on a scale from 0 to 1. If $|\alpha| < 1$, $\alpha \neq 0$, and using the following representation of the arctan function

$$\arctan x = \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} x^{2n+1}, \quad |x| < 1.$$

then the Gini index, which is defined as

$$G = 1 - 2\int_0^1 L_\alpha(p)\,\mathrm{d}p, \tag{14}$$

can be written as

$$G = -1 + \frac{2}{\arctan\alpha}\sum_{n=0}^{\infty}\frac{(-1)^n\alpha^{2n+1}}{2n+1}\int_0^1 (1-L_0(p))^{2n+1}\,\mathrm{d}p, \quad |\alpha|<1,\ \alpha\neq 0.$$

When $|\alpha| > 1, \alpha \neq 0$, more algebra is required, as we wish to obtain a closed form for the Gini index. In this case, and when the inverse of the initial Lorenz curve can be obtained simply, the Gini index is derived from the following result

**Proposition 3** *The Gini index for the Lorenz curve in (3) is given by*

$$G = \frac{2}{\arctan\alpha}\int_0^{\arctan\alpha} L_0^{-1}\left(1-\frac{1}{\alpha}\tan y\right)\mathrm{d}y - 1, \tag{15}$$

*for* $-\infty < \alpha < \infty$, $\alpha \neq 0$. *Here,* tan *is the circular tangent function and* $L_0^{-1}(\cdot)$ *is the inverse of the initial Lorenz curve.*

*Proof* By computing the inverse function of the Lorenz curve in (3) and using a result given by Anderson (1970), we have

$$\int_0^1 L_\alpha(p)\,\mathrm{d}p = 1 - \int_0^1 L_\alpha^{-1}(y)\,\mathrm{d}y.$$

Now, by computing the inverse of the Lorenz curve $L_\alpha(p)$, we obtain the result after some simple algebra. □

Expression (15) facilitates calculation of the Gini index, instead of using expression (14), especially when the inverse of the initial Lorenz curve can be computed straightforwardly.

For example, if we assume that the initial Lorenz curve is the egalitarian Lorenz curve then, by using (15), the Gini index is given by

$$G = 1 - \frac{\log(1+\alpha^2)}{\alpha\arctan\alpha}.$$

This result can also be obtained by performing integration by parts, taking into account that

$$\int_0^1 \arctan(\alpha(1-p))\,\mathrm{d}p = \frac{1}{\alpha}\arctan\alpha - \frac{1}{2\alpha^2}\log(1+\alpha^2).$$

An important generalization of the Gini index was proposed by Yitzhaki (1983), who suggested the generalized Gini index, which is defined as

$$G_\nu = 1 - \nu(\nu-1)\int_0^1 (1-p)^{\nu-2}L(p)\mathrm{d}p,$$

where $\nu > 1$ and $L(p)$ is the Lorenz curve. Of course, if $\nu = 2$, we obtain the Gini index. When $L_0(p) = p$, after some algebra, we obtain that the Yitzhaki index is given (see "Appendix") by

$$G_\nu = 1 - \frac{\alpha}{\arctan\alpha}\left[1 + \frac{\nu\alpha^2}{\nu+2}\,{}_2F_1\left(1, 1+\nu/2; 2+\nu/2, -\alpha^2\right)\right].$$

In the case of the Aggarwal and Singh initial Lorenz curve, using (15), the Gini index is given by

$$G = \frac{2\theta}{\arctan\alpha}\int_0^{\arctan\alpha}\frac{\alpha - \tan y}{\alpha\theta - \tan y}\,\mathrm{d}y - 1.$$

Then, the Gini index (see "Appendix") is expressed as

$$G = 2\theta\left[1 + \frac{\alpha(1-\theta)}{(1+\alpha^2\theta^2)\arctan\alpha}\left(\log\left(\frac{\theta\sqrt{1+\alpha^2}}{\theta - 1}\right) - \alpha\theta\arctan\alpha\right)\right] - 1.$$

Finally, assume the classical Pareto Lorenz curve as the initial Lorenz curve, and again using (15), the Gini index is given by

$$G = \frac{2}{\arctan\alpha}\int_0^{\arctan\alpha}\left[1 - \left(\frac{\tan y}{\alpha}\right)^{1/k}\right]\mathrm{d}y - 1.$$

The above integral is developed in the "Appendix," and the Gini index is found to be

$$G = 1 - \frac{2\alpha k}{(1+k)\arctan\alpha}\,{}_2F_1\left(1, \frac{1+k}{2k}; \frac{3k+1}{2}, -\alpha^2\right).$$

Using numerical integration techniques, Gini and Yitzhaki indices can also be calculated when other Lorenz curves are assumed as $L_0(p)$.

### 3.2 Pietra index

An interesting but less well-known index of inequality is the Pietra index, given by the proportion of total income that would need to be reallocated across the population to achieve perfect equality in income. This proportion is given by

$$P = \max_{0\le p\le 1}[p - L_\alpha(p)] = \frac{1}{2\mu}E|X - \mu|$$

and corresponds to the maximal vertical deviation between the Lorenz curve and the egalitarian line (Pietra 1915; Frosini 2012 calls this same index Pietra–Ricci index, owing to the extensive study made by Ricci (1916) on the same subject). Frosini (2005) also provides a simple graphical representation of this index.

Differentiating $p - L_\alpha(p)$ and using (3), we find that the Pietra index is attained for a value of $p$ satisfying the equation

$$\left[1+\alpha^2\left(1-L_0(p)\right)^2\right]\arctan\alpha - \alpha L_0'(p) = 0.$$

In particular, when $L_0(p) = p$, the maximum is attained when

$$p = 1 - \frac{1}{\alpha}\sqrt{\frac{\alpha-\arctan\alpha}{\arctan\alpha}}.$$

Then, the Pietra index is given, in this case, by

$$P = \frac{\arctan\left(\alpha\left(1-\frac{1}{\alpha}\sqrt{\frac{\alpha-\arctan\alpha}{\arctan\alpha}}\right)\right)}{\arctan\alpha} - \frac{1}{\alpha}\sqrt{\frac{\alpha-\arctan\alpha}{\arctan\alpha}}.$$

When the initial Lorenz curve considered is the Aggarwal and Singh Lorenz curve, the maximum is attained when

$$\begin{aligned} p_0 = \frac{1}{\theta^2+\alpha^2(1+\theta)^2}&\left[(1+\theta)\left(\alpha^2+\theta\left(1+\alpha^2\right)\right)\right.\\ &\left.-\frac{1}{\sqrt{\arctan\alpha}}\sqrt{\alpha(1+\theta)\left(\theta^2+\alpha^2(1+\theta)^2-\alpha(1+\theta)\arctan\alpha\right)}\right], \end{aligned}$$

and the Pietra index is then

$$P = p_0 + \frac{\arctan(\alpha(1-p_0(1-\theta)/(p_0-\theta)))}{\arctan\alpha} - 1.$$

Finally, for the Chotikapanich Lorenz curve, the Pietra index is

$$P = p_0 - 1 + \frac{1}{\arctan\alpha}\arctan\left[\alpha\left(1-\frac{e^{\theta p_0}-1}{e^\theta-1}\right)\right],$$

where $p_0$ is derived from

$$\begin{aligned} e^{\theta p_0} = \frac{1}{2\alpha\arctan\alpha}&\left[(\theta+2\alpha\arctan\alpha)e^\theta - 1\right.\\ &\left.-\sqrt{\left(\theta^2-(\arctan\alpha)^2\right)\left(e^\theta-1\right)^2+4\alpha\theta e^\theta\left(e^\theta-1\right)\arctan\alpha}\right]. \end{aligned}$$

Numerical computation can be used to obtain the Pietra index in other cases, when the initial Lorenz curve assumed is other than the egalitarian and Aggarwal and Singh Lorenz curves.

### 3.3 Population functions

In some particular cases, closed-form expressions can be obtained for the distribution functions. For example, if we assume that $L_0(p) = p$ we have, if $\alpha < 0$

$$F(x) = 1 + \frac{1}{\alpha}\kappa_1(x;\mu,\alpha), \quad \kappa_2(\mu,\alpha) \leq x \leq (1+\alpha^2)\kappa_2(\mu,\alpha)$$

and

$$F(x) = 1 - \frac{1}{\alpha}\kappa_1(x;\mu,\alpha), \quad \kappa_2(\mu,\alpha) \leq x \leq (1+\alpha^2)\kappa_2(\mu,\alpha),$$

if $\alpha > 0$, where $\kappa_1(x;\mu,\alpha) = \sqrt{\frac{\mu\alpha}{x \arctan\alpha} - 1}$ and $\kappa_2(\mu,\alpha) = \frac{\mu\alpha}{(1+\alpha^2)\arctan\alpha}$. The corresponding probability density functions are

$$f(x) = \frac{(1+\alpha^2)\kappa_2(\mu,\alpha)}{2\alpha x^2 \kappa_1(x;\mu,\alpha)}, \quad \kappa_2(\mu,\alpha) \leq x \leq (1+\alpha^2)\kappa_2(\mu,\alpha)$$

and

$$f(x) = -\frac{(1+\alpha^2)\kappa_2(\mu,\alpha)}{2\alpha x^2 \kappa_1(x;\mu,\alpha)}, \quad \kappa_2(\mu,\alpha) \leq x \leq (1+\alpha^2)\kappa_2(\mu,\alpha),$$

for $\alpha > 0$ and $\alpha < 0$, respectively.

Let $L_0(p)$ be the Aggarwal and Singh Lorenz curve. In this case, if $\alpha < 0$

$$F(x) = \kappa_1(\alpha,\theta) + \frac{1}{\sqrt{x}}\kappa_2(\mu,\alpha,\theta), \quad \left(\frac{\kappa_2(\mu,\alpha,\theta)}{\kappa_1(\alpha,\theta)}\right)^2 \leq x \leq \left(\frac{\kappa_2(\mu,\alpha,\theta)}{\kappa_1(\alpha,\theta)-1}\right)^2$$

and

$$F(x) = \kappa_1(\alpha,\theta) - \frac{1}{\sqrt{x}}\kappa_2(\mu,\alpha,\theta), \quad \left(\frac{\kappa_2(\mu,\alpha,\theta)}{\kappa_1(\alpha,\theta)}\right)^2 \leq x \leq \left(\frac{\kappa_2(\mu,\alpha,\theta)}{\kappa_1(\alpha,\theta)-1}\right)^2$$

if $\alpha > 0$, where

$$\kappa_1(\alpha,\theta) = \frac{(1+\theta)\left(\theta + \alpha^2(1+\theta)\right)}{\left(\theta^2 + \alpha^2(1+\theta)^2\right)},$$

$$\kappa_2(\mu,\alpha,\theta) = \frac{\sqrt{\alpha \arctan\alpha \left(\theta\mu\left(\theta^2 + \alpha^2(1+\theta)^2\right) - \alpha(1+\theta)^2 \arctan\alpha\right)}}{\left(\theta^2 + \alpha^2(1+\theta)^2\right)\arctan\alpha}.$$

The corresponding probability density functions are

$$f(x) = -\frac{1}{2x\sqrt{x}}\kappa_2(\mu,\alpha,\theta), \quad \left(\frac{\kappa_2(\mu,\alpha,\theta)}{\kappa_1(\alpha,\theta)}\right)^2 \leq x \leq \left(\frac{\kappa_2(\mu,\alpha,\theta)}{\kappa_1(\alpha,\theta)-1}\right)^2$$

and

$$f(x)=\frac{1}{2x\sqrt{x}}\kappa_2(\mu,\alpha,\theta),\quad \left(\frac{\kappa_2(\mu,\alpha,\theta)}{\kappa_1(\alpha,\theta)}\right)^2\leq x\leq\left(\frac{\kappa_2(\mu,\alpha,\theta)}{\kappa_1(\alpha,\theta)-1}\right)^2$$

for $\alpha>0$ and $\alpha<0$, respectively.

Finally, for the arctan Chotikapanich Lorenz curve, the population function becomes

$$F(x)=\frac{1}{\theta}\log\left[\frac{\theta\mu(e^\theta-1)+2\alpha xe^\theta\arctan\alpha-\sqrt{e^\theta-1}H(\alpha,\theta,\mu,x)}{2x\alpha\arctan\alpha}\right],$$

where

$$H(\alpha,\theta,\mu,x)=\theta^2\mu^2(e^\theta-1)+4x\arctan\alpha\left[x\arctan\alpha+e^\theta(\alpha\theta\mu-x\arctan\alpha)\right],$$

begin $-\infty<\alpha<\infty,\alpha\neq 0$ and

$$\frac{\alpha\theta\mu}{(\alpha^2+1)\left(e^\theta-1\right)\arctan\alpha}\leq x\leq\frac{\alpha\theta\mu}{(\alpha^2+1)\left(e^\theta-1\right)\arctan\alpha}.$$

## 4 Numerical application

To compare the performance of the functional forms given in (10), (11) and (12), we used the US data (for 2009 and 2013) obtained from the US Census Bureau, Current Population Survey, 2014 Annual Social and Economic Supplement (see "Appendix, Tables 5 and 6"). Three methods of estimation are considered, as described below.

### 4.1 Nonlinear least squares estimators

These are defined by the estimators which minimize the sum of the squared differences between the predicted and observed values. For a particular Lorenz curve $L_\alpha(p)$, the minimization is associated with the expression

$$\sum_{i=1}^{n}(p_i-L_\alpha(p_i))^2,$$

where the points $(p_i,L_\alpha(p_i))_{i=1}^n$ are available from an empirical Lorenz curve.

From the approximation given in (7), we consider as initial estimates those obtained by least squares, replacing $L_0(p)$ for the classical expression and in every case mapping from the observations to the estimated parameters. This expression can also be employed to obtain estimates by the method proposed by Castillo et al. (1998). In this case, we begin by considering a single point $(p_i,q_i)$ of the empirical Lorenz curve, and by substituting in (7), we obtain the simple estimate for $\alpha$ given by

**Table 1** Results for the parameter estimates and MSE and MAX criteria

| Model | Estimated parameters | | SSE[a] | MAX[b] | Estimated indices | |
|---|---|---|---|---|---|---|
| | $\widehat{\alpha}$ | $\widehat{\theta}$ | | | Gini | Pietra |
| *Based on 2003 data for the USA* | | | | | | |
| Classical | | | | | | |
| Aggarwal and Singh | | 2.98072 | 0.0104026 | 0.0463689 | 0.433071 | 0.332259 |
| Pareto | | 0.40817 | 0.0782100 | 0.0830288 | 0.420283 | 0.319010 |
| Chotikapanich | | 3.03685 | 0.0163449 | 0.0627102 | 0.442232 | 0.339143 |
| Arctan | | | | | | |
| Egalitarian | 3.60671 | | 0.0207202 | 0.0731232 | 0.437163 | 0.343442 |
| Aggarwal and Singh | 0.25427 | 2.84932 | 0.0103904 | 0.0472890 | 0.433214 | 0.332529 |
| Pareto | 2.26125 | 0.77681 | 0.0091147 | 0.0212034 | 0.435615 | 0.330779 |
| Chotikapanich | 0.91206 | 2.27681 | 0.0156580 | 0.0635445 | 0.439873 | 0.339427 |
| *Based on 2013 data for the USA* | | | | | | |
| Classical | | | | | | |
| Aggarwal and Singh | | 3.29329 | 0.0128622 | 0.0430045 | 0.453747 | 0.348964 |
| Pareto | | 0.44192 | 0.0823078 | 0.0752993 | 0.387038 | 0.292319 |
| Chotikapanich | | 3.37974 | 0.0451976 | 0.0595370 | 0.478752 | 0.368801 |
| Arctan | | | | | | |
| Egalitarian | 4.17665 | | 0.0559155 | 0.0658484 | 0.477561 | 0.376789 |
| Aggarwal and Singh | $6.7 \times 10^{-8}$ | 3.29329 | 0.0128622 | 0.0430045 | 0.453747 | 0.348964 |
| Pareto | 2.03291 | 0.71972 | 0.0042508 | 0.0213570 | 0.444280 | 0.335693 |
| Chotikapanich | 1.16084 | 2.28120 | 0.0426761 | 0.0601549 | 0.475442 | 0.369078 |

[a] SSE denotes the sum of squares of model estimation errors
[b] MAX denotes the maximum absolute error

$$\widehat{\alpha}_i \approx \frac{L_0(p_i;\widehat{\phi}) - q_i}{q_i(1 - L_0(p_i;\widehat{\phi}))}, \quad i = 1, 2, \ldots, n, \tag{16}$$

where $\widehat{\phi}$ is the least squares estimate obtained from the classical Lorenz curve, which depends on parameter $\phi$ (which is a vector of parameters when the classical Lorenz curve depends on more than one parameter). By combining all the initial estimators (16) using a function such as the mean or median, the final estimators are obtained. For example, if we use the mean function, the final estimation of $\alpha$ will be $\widehat{\alpha} = \frac{1}{n}\sum_{i=1}^{n} \widehat{\alpha}_i$.

Finally, the results for the two data sets, 2009 and 2013, are shown in Table 1. The parameter estimates, the mean squared error (MSE) and the maximum absolute error (MAX) were computed for the two data sets considered. The corresponding table shows that the new models provide better results in terms of smaller MSE, MAX, Gini and Pietra indices (the empirical Gini, computed according to Brown's formula, and Pietra indices give the results 0.450007 and 0.324733, respectively, for the 2003 data and 0.457607 and 0.330401 for the 2013 data) with respect to the initial Lorenz curves considered, and that the best fit is obtained with the new functional forms proposed.

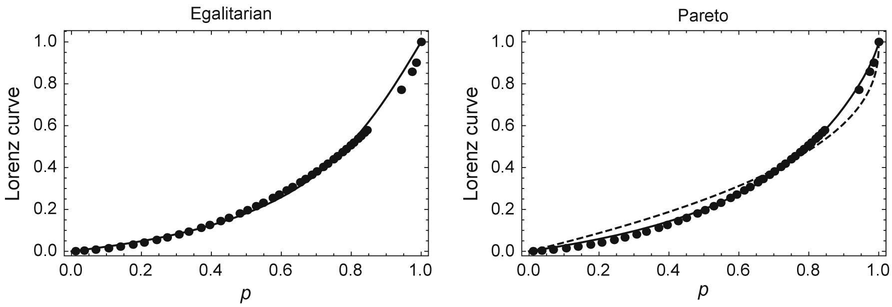


**Fig. 1** Lorenz curves for 2003 US income data based on nonlinear least square estimates. *Dashed curves* represent the classical model and *continuous curves*, the arctan model

Figure 1 presents a graphical comparison between the empirical Lorenz curves and the corresponding estimated Lorenz curves based on the nonlinear least squares estimators for the Egalitarian and Pareto cases.

### 4.2 Maximum likelihood estimation based on the use of the population function

Maximum likelihood estimation based on the use of the population function was also studied, using the cumulative distribution functions given in Sect. 3.3. When data are grouped, let $n_i$ be the number of observations in the interval $(c_{j-1}, c_j]$. The log-likelihood function is then,

$$\ell(\phi) = \sum_{i=1}^{n} n_i \log \left[ F(x_i|\phi) - F(x_{i-1}|\phi \right],$$

where $n$ is the sample size and $\phi$ the parameter/s to be estimated. See Chotikapanich (2008) for details. From Table 2, we can see that the arctan model provides the value of the maximum of the log-likelihood function in a better way than does the Dirichlet distribution.

Because there is a mapping from the Lorenz curve to the density of the data and in order to correct standard errors for model misspecification, we have estimated the parameters of interest by maximizing the log-likelihood and obtained robust (sandwich) standard errors. See Freedman (2006) for details.

Finally, when the population function associated with a given Lorenz curve is not known, estimation based on the use of the Dirichlet distribution is adequate for comparing different models (see Chotikapanich and Griffiths 2002).

### 4.3 Model validation

For the situation in which the models are non-nested, a Vuong test was conducted to compare the estimates of the different Lorenz curves. In this regard, we test the null hypothesis that the two models are equally close to the actual model, against the alternative that one of them is closer (Vuong 1989). The $z$-statistic is

**Table 2** MLE based on cumulative distribution function

| Model | Estimated parameter | | $\ell_{\max}$ | Gini | Pietra |
|---|---|---|---|---|---|
| | $\widehat{\alpha}$ | $\widehat{\theta}$ | | | |
| *Based on 2003 data for the USA* | | | | | |
| Classical | | | | | |
| Aggarwal and Singh | | 3.20946 (0.457267) | −364122 | 0.448395 | 0.344627 |
| Pareto | | 0.351708 (0.055653) | −474591 | 0.479609 | 0.367763 |
| Chotikapanich | | 4.12416 (0.364762) | −420247 | 0.547939 | 0.426463 |
| Arctan | | | | | |
| Egalitarian | 5.84025 (0.334026) | | −415677 | 0.565166 | 0.450949 |
| Aggarwal and Singh | $5.08 \times 10^{-8}(2.02 \times 10^{-8})$ | 3.20946 (0.457267) | −364122 | 0.448395 | 0.344627 |
| Chotikapanich | 1.343 (2.00614) | 2.84309 (1.165930) | −406952 | 0.550498 | 0.435529 |
| *Based on 2013 data for the USA* | | | | | |
| Classical | | | | | |
| Aggarwal and Singh | | 2.43183 (0.631985) | −359208 | 0.391275 | 0.298866 |
| Pareto | | 0.420322 (0.125857) | −472147 | 0.408131 | 0.309206 |
| Chotikapanich | | 3.28319 (0.554086) | −410022 | 0.468776 | 0.360652 |
| Arctan | | | | | |
| Egalitarian | 4.57089 (0.465012) | | −407555 | 0.501865 | 0.397088 |
| Aggarwal and Singh | $4.05 \times 10^{-10}(3.75 \times 10^{-10})$ | 2.43183 (0.631985) | −359208 | 0.391275 | 0.298866 |
| Chotikapanich | 2.25203 (2.443510) | 1.238 (1.547920) | −403039 | 0.483135 | 0.378491 |

The robust standard errors are shown in parentheses

$$Z = \frac{1}{\omega\sqrt{n}}\left(\ell\left(\widehat{\theta_1}\right) - \ell\left(\widehat{\theta_2}\right)\right),$$

where $\widehat{\theta_1}$ and $\widehat{\theta_2}$ are vectors of the estimated parameters and

$$\omega^2 = \frac{1}{n}\sum_{i=1}^{n}\left[\log\left(\frac{f\left(x_i|\widehat{\theta_1}\right)}{g\left(x_i|\widehat{\theta_2}\right)}\right)\right]^2 - \left[\frac{1}{n}\sum_{i=1}^{n}\log\left(\frac{f\left(x_i|\widehat{\theta_1}\right)}{g\left(x_i|\widehat{\theta_2}\right)}\right)\right]^2$$

where $f$ and $g$ represent the probability density functions of the two models to be compared, respectively.

Due to the asymptotically normal behavior of the $Z$ statistic, the null hypothesis is rejected in favor of the alternative that $f$ occurs with a significance level $\alpha$, when $Z > z_{1-\alpha}$, where $z_{1-\alpha}$ is the $(1-\alpha)$ quantile of the standard normal distribution.

To work with this test, we choose a critical value from the standard normal distribution that corresponds to the desired level of significance (e.g., for $c = 1.96$; $\Pr(z \geq |\pm c|) = 0.05$). Then, if $z > c$, we reject the null hypothesis that the models

**Table 3** Vuong test comparison of non-nested models

| | Aggarwal and Singh | Pareto | Chotikapanich |
|---|---|---|---|
| Arctan Egalitarian | −9.2821 | 8.86773 | 0.770538 |
| | −5.2335 | 5.7230 | 0.361578 |
| Arctan Aggarwal and Singh | – | 9.82886 | 5.60386 |
| | | 5.69444 | 3.63771 |
| Arctan Chotikapanich | −8.58913 | 6.017 | – |
| | −4.75552 | 3.48425 | – |

2003 year above and 2013 below

**Table 4** Log-likelihood ratio comparison of nested models

| Model | 2003 | 2013 |
|---|---|---|
| Aggarwal and Singh | 9828 | – |
| Chotikapanich | 26590 | 13966 |

are the same, in favor of the alternative that $f$ is better than $g$. Thus, if $z < c$, we reject the null hypothesis that the models are the same in favor of the alternative that $g$ is better than $f$, while if $z \leq c$, we cannot reject the null hypothesis that the models are the same. Under this criterion, and from Table 3, we conclude that the classical Aggarwal and Singh Lorenz curve performs all the arctan models proposed and that the Chotikapanich Lorenz curve performs the arctan Egalitarian Lorenz curve. In the remaining cases, the arctan models are better than the Pareto and the Chotikapanich Lorenz curves.

Finally, we examined whether likelihood ratio tests suggested that nested versions were adequate. This test was computed, and the results obtained are shown in Table 4. As we can see, the arctan model performs the classical model.

## 5 Conclusions

The proposed family of Lorenz curves seems to be a worthy addition to the existing class of single parameter Lorenz curves. The family was applied to two data sets with satisfactory results, using least squares and maximum likelihood. Thus, the new specification is well capable of modeling income data.

**Acknowledgments** The authors thank the two anonymous referees and the Associate Editor for their valuable comments and suggestions. EGD was partially funded by Grant ECO2013-47092 (Ministerio de Economía y Competitividad, Spain).

## Appendix

To compute the integral

$$G_\nu = 1 - \frac{\nu(\nu-1)}{\arctan\alpha}\int_0^1 (1-p)^{\nu-2}\arctan(\alpha(1-p))\,\mathrm{d}p$$

we perform integration by parts, which gives

$$G_\nu = 1 + \nu + \frac{\alpha\nu}{\arctan\alpha}\int_0^1 \frac{(1-p)^{\nu-1}}{1+\alpha^2(1-p)^2}\,\mathrm{d}p.$$

The above integral is obtained by making the change of variable $\omega = 1 - p$, and the result is obtained after some algebra, taking into account (6).

To obtain the integral

$$\int_0^{\arctan\alpha} \frac{\alpha - \tan y}{\alpha\theta - \tan y}\,\mathrm{d}y$$

we make the change of variable $\omega = \alpha\theta - \tan y$ and thus obtain the rational integral

$$\int_0^{\arctan\alpha} \frac{\alpha - \tan y}{\alpha\theta - \tan y}\,\mathrm{d}y = -\int_{\alpha\theta}^{\alpha(\theta-1)} \frac{\alpha(1-\theta)+\omega}{\omega(1+(\alpha\theta-\omega)^2}\,\mathrm{d}\omega,$$

which is simple to calculate.

In order to obtain the integral $\int_0^{\arctan\alpha}(\tan y)^{1/k}\,\mathrm{d}y$ we make the change of variable $\omega = \frac{1}{\alpha^2}\tan^2 y$, giving the integral

$$\frac{\alpha^{1/k+1}}{2}\int_0^1 \omega^{\frac{1}{2}\left(\frac{1}{k}-1\right)}(1+\alpha^2\omega)^{-1}\,\mathrm{d}\omega.$$

From which the result is obtained after some algebra, taking into account (6).

**Table 5** Distribution of income to \$250,000 or more for households

| Income of household | | Mean income | Number | Cumulative freq. | |
|---|---|---|---|---|---|
| Lower | Upper | | | Population | Income |
| 2500 | 4999 | 3810 | 1266 | 0.011565 | 0.000729 |
| 5000 | 7499 | 6394 | 2717 | 0.036386 | 0.003355 |
| 7500 | 9999 | 8722 | 3593 | 0.069210 | 0.008093 |
| 10,000 | 12,499 | 11,174 | 4032 | 0.106044 | 0.014904 |
| 12,500 | 14,999 | 13,636 | 3708 | 0.139918 | 0.022548 |
| 15,000 | 17,499 | 16,104 | 3933 | 0.175848 | 0.032123 |
| 17,500 | 19,999 | 18,653 | 3501 | 0.207831 | 0.041995 |
| 20,000 | 22,499 | 21,073 | 3888 | 0.243349 | 0.054381 |
| 22,500 | 24,999 | 23,720 | 3327 | 0.273743 | 0.066312 |
| 25,000 | 27,499 | 26,067 | 3683 | 0.307389 | 0.080825 |
| 27,500 | 29,999 | 28,667 | 3035 | 0.335115 | 0.093978 |
| 30,000 | 32,499 | 31,019 | 3882 | 0.370578 | 0.112183 |
| 32,500 | 34,999 | 33,655 | 2677 | 0.395034 | 0.125803 |

**Table 5** continued

| Income of household | | Mean income | Number | Cumulative freq. | |
|---|---|---|---|---|---|
| Lower | Upper | | | Population | Income |
| 35,000 | 37,499 | 36,058 | 3417 | 0.426250 | 0.144430 |
| 37,500 | 39,999 | 38,643 | 2607 | 0.450066 | 0.159660 |
| 40,000 | 42,499 | 41,009 | 3420 | 0.481309 | 0.180862 |
| 42,500 | 44,999 | 43,644 | 2381 | 0.503060 | 0.196572 |
| 45,000 | 47,499 | 46,024 | 2721 | 0.527918 | 0.215504 |
| 47,500 | 49,999 | 48,630 | 2227 | 0.548262 | 0.231876 |
| 50,000 | 52,499 | 50,971 | 3012 | 0.575778 | 0.255086 |
| 52,500 | 54,999 | 53,682 | 1963 | 0.593711 | 0.271016 |
| 55,000 | 57,499 | 56,065 | 2343 | 0.615115 | 0.290875 |
| 57,500 | 59,999 | 58,651 | 1833 | 0.631861 | 0.307128 |
| 60,000 | 62,499 | 60,978 | 2442 | 0.654169 | 0.329639 |
| 62,500 | 64,999 | 63,664 | 1648 | 0.669225 | 0.345500 |
| 65,000 | 67,499 | 66,004 | 2007 | 0.687559 | 0.365527 |
| 67,500 | 69,999 | 68,641 | 1550 | 0.701719 | 0.381611 |
| 70,000 | 72,499 | 71,007 | 1995 | 0.719944 | 0.403027 |
| 72,500 | 74,999 | 73,665 | 1398 | 0.732716 | 0.418595 |
| 75,000 | 77,499 | 75,998 | 1835 | 0.749479 | 0.439678 |
| 77,500 | 79,999 | 78,646 | 1243 | 0.760835 | 0.454457 |
| 80,000 | 82,499 | 80,969 | 1575 | 0.775223 | 0.473736 |
| 82,500 | 84,999 | 83,688 | 1162 | 0.785838 | 0.488437 |
| 85,000 | 87,499 | 86,062 | 1362 | 0.798281 | 0.506157 |
| 87,500 | 89,999 | 88,687 | 968 | 0.807124 | 0.519136 |
| 90,000 | 92,499 | 91,029 | 1362 | 0.819566 | 0.537879 |
| 92,500 | 94,999 | 93,674 | 888 | 0.827679 | 0.550454 |
| 95,000 | 97,499 | 96,117 | 1114 | 0.837855 | 0.566641 |
| 97,500 | 99,999 | 98,677 | 804 | 0.845200 | 0.578635 |
| 100,000 | 149,999 | 118,880 | 10,719 | 0.943123 | 0.771276 |
| 150,000 | 199,999 | 168,806 | 3372 | 0.973928 | 0.857328 |
| 200,000 | 249,999 | 219,390 | 1307 | 0.985867 | 0.900677 |
| ≥250,000 | | 424,693 | 1547 | 1.000000 | 1.000000 |

USA 2003. *Source*: US Census Bureau, Current Population Survey, 2014 Annual Social and Economic Supplement

**Table 6** Distribution of income to $250,000 or more for households

| Income of household | | Mean income | Number | Cumulative freq. | |
|---|---|---|---|---|---|
| Lower | Upper | | | Population | Income |
| 5000 | 9999 | 7983 | 4859 | 0.040875 | 0.004345 |
| 10,000 | 14,999 | 12,425 | 6693 | 0.097180 | 0.013661 |
| 15,000 | 19,999 | 17,249 | 7321 | 0.158767 | 0.027808 |

**Table 6** continued

| Income of household | | Mean income | Number | Cumulative freq. | |
|---|---|---|---|---|---|
| Lower | Upper | | | Population | Income |
| 20,000 | 24,999 | 22,254 | 6577 | 0.214096 | 0.044205 |
| 25,000 | 29,999 | 27,164 | 6302 | 0.267111 | 0.063383 |
| 30,000 | 34,999 | 32,057 | 6454 | 0.321405 | 0.086561 |
| 35,000 | 39,999 | 37,111 | 5827 | 0.370424 | 0.110786 |
| 40,000 | 44,999 | 42,035 | 5565 | 0.417239 | 0.136992 |
| 45,000 | 49,999 | 47,057 | 5286 | 0.461707 | 0.164858 |
| 50,000 | 54,999 | 51,940 | 5198 | 0.505434 | 0.195104 |
| 55,000 | 59,999 | 57,102 | 4349 | 0.542020 | 0.222924 |
| 60,000 | 64,999 | 61,914 | 4422 | 0.579220 | 0.253596 |
| 65,000 | 69,999 | 67,049 | 3818 | 0.611338 | 0.282274 |
| 70,000 | 74,999 | 72,012 | 3872 | 0.643911 | 0.313511 |
| 75,000 | 79,999 | 77,012 | 3702 | 0.675054 | 0.345449 |
| 80,000 | 84,999 | 82,054 | 3384 | 0.703521 | 0.376556 |
| 85,000 | 89,999 | 87,038 | 2622 | 0.725579 | 0.402122 |
| 90,000 | 94,999 | 92,100 | 2691 | 0.748217 | 0.429887 |
| 95,000 | 99,999 | 97,069 | 2288 | 0.767464 | 0.454768 |
| 100,000 | 104,999 | 101,891 | 2563 | 0.789025 | 0.484023 |
| 105,000 | 109,999 | 107,100 | 1922 | 0.805194 | 0.507084 |
| 110,000 | 114,999 | 111,993 | 1848 | 0.820740 | 0.530269 |
| 115,000 | 119,999 | 117,111 | 1528 | 0.833594 | 0.550316 |
| 120,000 | 124,999 | 121,889 | 1599 | 0.847046 | 0.572150 |
| 125,000 | 129,999 | 127,065 | 1389 | 0.858730 | 0.591922 |
| 130,000 | 134,999 | 132,061 | 1291 | 0.869591 | 0.611022 |
| 135,000 | 139,999 | 137,174 | 1082 | 0.878693 | 0.627649 |
| 140,000 | 144,999 | 141,980 | 1081 | 0.887787 | 0.644843 |
| 145,000 | 149,999 | 146,949 | 963 | 0.895888 | 0.660696 |
| 150,000 | 154,999 | 151,594 | 1298 | 0.906807 | 0.682740 |
| 155,000 | 159,999 | 157,250 | 847 | 0.913933 | 0.697661 |
| 160,000 | 164,999 | 161,893 | 771 | 0.920419 | 0.711644 |
| 165,000 | 169,999 | 167,100 | 636 | 0.925769 | 0.723550 |
| 170,000 | 174,999 | 171,822 | 623 | 0.931010 | 0.735542 |
| 175,000 | 179,999 | 177,134 | 546 | 0.935603 | 0.746376 |
| 180,000 | 184,999 | 181,670 | 529 | 0.940053 | 0.757143 |
| 185,000 | 189,999 | 186,844 | 427 | 0.943645 | 0.766080 |
| 190,000 | 194,999 | 192,136 | 428 | 0.947246 | 0.775293 |
| 195,000 | 199,999 | 197,196 | 358 | 0.950257 | 0.783202 |

**Table 6** continued

| Income of household | | Mean income | Number | Cumulative freq. | |
|---|---|---|---|---|---|
| Lower | Upper | | | Population | Income |
| 200,000 | 249,999 | 220,406 | 2600 | 0.972130 | 0.847399 |
| ≥250,000 | | 411,160 | 3313 | 1.000000 | 1.000000 |

USA 2013. *Source*: US Census Bureau, Current Population Survey, 2014 Annual Social and Economic Supplement

## References


Aggarwal V, Singh R (1984) On optimum stratification with proportional allocation for a class of Pareto distributions. Commun Stat Theory Methods 13:3017–3116

Anderson N (1970) Integration of inverse functions. Math Gaz 54(387):52–53

Arnold BC (1986) A class of hyperbolic Lorenz curves. Sankhyā: Indian J Stat, Ser B 48(3):427–436

Basmann RL, Hayes KL, Slottje DJ, Johnson JD (1990) A general functional form for approximating the Lorenz curve. J Econom 43:77–90

Burrell QL (1992) The Gini index and the Leimkuhler curve for bibliometric processes. Inf Process Manag 28:19–33

Burrell QL (2005) Symmetry and other transformation features of Lorenz/Leimkuhler representations of informetric data. Inf Process Manag 41:1317–1329

Castellanos D (1988) The ubiquitous pi. Math Mag 61:67–98

Castillo E, Hadi AS, Sarabia JM (1998) A method for estimating Lorenz curves. Commun Stat-Theory Methods 27:2037–2063

Chotikapanich D (1993) A comparison of alternative functional forms for the Lorenz curve. Econ Lett 41:129–138

Chotikapanich D (2008) Modeling income distributions and Lorenz curves. Springer, Berlin

Chotikapanich D, Griffiths WE (2002) Estimating Lorenz curves using a Dirichlet distribution. J Bus Econ Stat 20(2):290–295

Freedman DA (2006) On the so-called "Huber sandwich estimator" and "robust standard errors". Am Stat 60(4):299–302

Frosini BV (2005) Inequality measures for histograms. Statistica 65:27–40

Frosini BV (2012) Approximation and decomposition of Gini, Pietra-index and Theil inequality measures. Empir Econ 43:175–197

Gastwirth JL (1971) A general definition of the Lorenz curve. Econometrica 39:1037–1039

Gupta MR (1984) Functional forms for fitting the Lorenz curve. Econometrica 52:1313–1314

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

Ogwang T, Rao URG (1996) A new functional form for approximating the Lorenz curve. Econ Lett 52:21–29

Ortega P, Martín G, Fernandez A, Ladoux M, García A (1991) A new functional form for estimating Lorenz curves. Rev Income Wealth 37:447–452

Pietra G (1915) Delle relazioni tra gli indici di variabilit. Atti Regio Istituto Veneto 74(II):775–792

Rajan S, Wang S, Inkol R, Joyal A (2006) Efficient approximations for the arctangent function. IEEE Signal Process Mag 23(3):108–111

Ricci U (1916) Lindice di variabilit e la curva dei redditi. Giornale degli economisti e Rivista di statistica 53:177–228

Ryu HK, Slottje DJ (1996) Two flexible functional form approaches for approximating the Lorenz curve. J Econom 72:251–274

Sarabia JM, Castillo E, Slottje DJ (1999) An ordered family of Lorenz curves. J Econom 91:43–60

Sarabia JM, Gómez-Déniz E, Sarabia M, Prieto F (2010) A general method for generating parametric Lorenz and Leimkuhler curves. J Informetr 4:424–539

Sarabia JM, Sarabia M (2008) Explicit expressions for the Leimkuhler curve in parametric families. Inf Process Manag 44:1808–1818

Vuong Q (1989) Likelihood ratio tests for model selection and non-nested hypotheses. Econometrica 57:307–333

Yitzhaki S (1983) On an extension of the Gini inequality index. Int Econ Rev 24:617–628