

Article

Lorenz Surfaces Based on the Sarmanov–Lee Distribution with Applications to Multidimensional Inequality in Well-Being

José María Sarabia ^{1,*},† and Vanesa Jorda ^{2,*},†

¹ Department of Quantitative Methods, CUNEF University, Leonardo Prieto Castro 2, 28040 Madrid, Spain

² Department of Economics, University of Cantabria, Avda de los Castros s/n, 39005 Santander, Spain

* Correspondence: josemaria.sarabia@cunef.edu (J.M.S.); jordav@unican.es (V.J.)

† These authors contributed equally to this work.

Received: 14 October 2020; Accepted: 17 November 2020; Published: 23 November 2020



Abstract: The purpose of this paper is to derive analytic expressions for the multivariate Lorenz surface for a relevant type of models based on the class of distributions with given marginals described by Sarmanov and Lee. The expression of the bivariate Lorenz surface can be conveniently interpreted as the convex linear combination of products of classical and concentrated univariate Lorenz curves. Thus, the generalized Gini index associated with this surface is expressed as a function of marginal Gini indices and concentration indices. This measure is additively decomposable in two factors, corresponding to inequality within and between variables. We present different parametric models using several marginal distributions including the classical Beta, the GB1, the Gamma, the lognormal distributions and others. We illustrate the use of these models to measure multidimensional inequality using data on two dimensions of well-being, wealth and health, in five developing countries.

Keywords: multivariate lorenz surface; Sarmanov–Lee distribution; generalized Gini index; well-being

1. Introduction

Even though a consensus has slowly emerged among scholars that economic inequality is a multidimensional construct, most of the early work in this area relies almost exclusively on income variables. In the last decades, however, academics' interest has shifted from income alone to a more comprehensive conception of well-being, which has an intrinsic multidimensional nature (see, e.g., [1–3]). Academics have repeatedly argued that the same monetary amount may yield different standards of quality of life for people with different needs [4,5]. Hence, the level of income is a crude proxy of well-being that an individual enjoys.

The renewed interest in well-being inequality has motivated the development of different approaches to measuring this phenomenon. Ideally, the evaluation of multidimensional inequality should consider both inequality within each dimension and the degree of association between dimensions. Multidimensional measures that consider both types of inequalities have been proposed by [6–13]. However, depending on their properties, different measures might reveal completely different results. When this happens, graphical methods can be a powerful tool to explore where the redistributive movements take place at different parts of the distribution. The extension of the univariate Lorenz curve (LC) to higher dimensions is therefore essential to further explore the distributional patterns of well-being. The three existing definitions of the multivariate LC were proposed by Taguchi [14,15], Arnold [16] and Koshevoy and Mosler [17], who introduced the concepts of Lorenz zonoid and Gini zonoid index.

In this paper, using the definition proposed by Arnold [16], we derive analytic expressions for the bivariate Lorenz surface for different formulations of the underlying bivariate distribution.

A preliminary account of this formulation is provided in [18]. In the present paper, we generalize these results and apply them to the study of multidimensional inequality in well-being. We study a relevant type of models based on the class of bivariate distributions with given marginals described by Sarmanov and Lee [19,20]. This model presents several advantages. The expression of the bivariate Lorenz surface can be easily expressed as a convex linear combination of classical and generalized LCs. The generalized Gini index is also available in a closed form and can be decomposed into two factors, corresponding to the equality within and between variables. We explore several models with different types of marginals. We also discuss the extension to higher dimensions.

In the following section, we present preliminary results including the definition of univariate Lorenz and concentration curves and a short review of the three different definitions of bivariate Lorenz surfaces proposed in the literature. In Section 3, we obtain a closed formula of the bivariate Sarmanov–Lee Lorenz surface and its corresponding generalized Gini index. The decomposition of this index into two factors is also derived. Thereafter, we explore different models of bivariate Lorenz surfaces in Section 4. We study the extension to higher dimensions in Section 5. We provide an application to measurement of multidimensional inequality in well-being in Section 6. Finally, some conclusions are sketched in Section 7.

2. Preliminary Results

2.1. Univariate Lorenz and Concentration Curves

We denote the class of univariate distributions functions with positive finite expectations by \mathcal{L} and denote by \mathcal{L}_+ the class of all distributions in \mathcal{L} with $F(0) = 0$ corresponding to non-negative random variables. We use the following definition by Gastwirth [21].

Definition 1. The Lorenz curve L of a random variable X with cumulative distribution function $F \in \mathcal{L}$ is,

$$L(u; F) = \frac{\int_0^u F^{-1}(y)dy}{\int_0^1 F^{-1}(y)dy} = \frac{\int_0^u F^{-1}(y)dy}{E(X)}, \quad 0 \leq u \leq 1, \tag{1}$$

where

$$\begin{aligned} F^{-1}(y) &= \sup\{x : F(x) \leq y\}, & 0 \leq y < 1, \\ &= \sup\{x : F(x) < 1\}, & y = 1, \end{aligned}$$

is the right continuous inverse distribution function or quantile function corresponding to F .

Now, we present the concept of concentration curve introduced in [22]. Let $g(x)$ be a continuous function of x such that its first derivative exists and $g(x) \geq 0$. If the mean $E_F[g(X)]$ exists, then one can define

$$L_g(y; F) = \frac{\int_0^x g(x)dF(x)}{E_F[g(X)]}$$

where $y = g(x)$ and $f(x)$ and $F(x)$ are, respectively, the probability density function (pdf) and the cumulative distribution function (cdf) of the random variable X . The implicit relation between $L_g(g(x); F)$ and $F(x)$ will be called the concentration curve of the function $g(X)$. If $\eta_g(x)$ and $\eta_{g^*}(x)$ denote the elasticities of $g(x)$ and $g^*(x)$, the concentration curve for the function $g(x)$ will lie above (below) the concentration curve for the function $g^*(x)$ if $\eta_g(x)$ is less (greater) than $\eta_{g^*}(x)$ for all $x \geq 0$. The concentration curve admits the simple implicit representation,

$$L_g(u; F) = \frac{1}{E_F[g(X)]} \int_0^u g[F^{-1}(t)]dt, \tag{2}$$

which will be used in the following sections.

2.2. The Three Alternative Definitions of Bivariate Lorenz Surface

In this section, we discuss briefly the three definitions proposed in the literature for constructing a bivariate Lorenz surface. These definitions were suggested by Taguchi [14,15], Arnold [16] and Koshevoy and Mosler [17]. We consider the bivariate definition, which can be extended to higher dimensions.

First, we present the Taguchi [14] definition of bivariate Lorenz surface, which was the first proposal.

Definition 2. The Lorenz surface of F_{12} is the set of points $(s, t, L(s, t))$ in \mathbb{R}_+^3 defined by,

$$s = \int_0^u \int_0^v f_{12}(x_1, x_2) dx_1 dx_2, \quad t = \frac{\int_0^u \int_0^v x_1 f_{12}(x_1, x_2) dx_1 dx_2}{E(X_1)},$$

$$L(s, t; F_{12}) = \frac{\int_0^u \int_0^v x_2 f_{12}(x_1, x_2) dx_1 dx_2}{E(X_2)}.$$

This definition was investigated by Taguchi [14,15,23]. The definition is not symmetric in (s, t) , and its extension to higher dimensions does not look simple.

The following definition was proposed by Arnold [16,24], which is an extension of (1) to higher dimensions. Let $\mathbf{X} = (X_1, X_2)^T$ be a bivariate random variable with bivariate probability distribution function F_{12} on \mathbb{R}_+^2 having finite second and positive first moments. We denote by $F_i, i = 1, 2$ the marginal cdf corresponding to $X_i, i = 1, 2$, respectively.

Definition 3. The Lorenz surface of F_{12} is the graph of the function,

$$L(u_1, u_2; F_{12}) = \frac{\int_0^{s_1} \int_0^{s_2} x_1 x_2 dF_{12}(x_1, x_2)}{\int_0^\infty \int_0^\infty x_1 x_2 dF_{12}(x_1, x_2)}, \tag{3}$$

where

$$u_1 = \int_0^{s_1} dF_1(x_1), \quad u_2 = \int_0^{s_2} dF_2(x_2), \quad 0 \leq u_1, u_2 \leq 1.$$

Now, we denote by F_a the one-point distribution at $a \in \mathbb{R}_+^2$, which corresponds to the bivariate egalitarian distribution at a . In consequence, the egalitarian distribution has bivariate Lorenz surface $L(u_1, u_2; F_a) = u_1 u_2$.

The two-attribute Gini index $G(F_{12})$ is defined as,

$$G(F_{12}) = 4 \int_0^1 \int_0^1 [u_1 u_2 - L(u_1, u_2; F_{12})] du_1 du_2 \tag{4}$$

where the egalitarian surface is given by $L(u_1, u_2; F_a) = u_1 u_2$. This index evaluates well-being inequality using a composite index of the form:

$$W(x_1, x_2) = x_1 \cdot x_2.$$

Given the multiplicative nature of the index, well-being dimensions are neither perfect substitutes nor complements. The marginal rate of substitution is $-x_2/x_1$.

Previous definition has not been explored in detail in the literature. We highlight some of its properties:

1. The marginal LCs can be obtained as $L(u_1; F_1) = \lim_{u_2 \rightarrow \infty} L(u_1, u_2; F_{12})$ and $L(u_2; F_2) = \lim_{u_1 \rightarrow \infty} L(u_1, u_2; F_{12})$.

2. The bivariate Lorenz surface does not depend on changes of scale in the marginals.
3. If F_{12} is a product distribution function, then

$$L(u_1, u_2; F_{12}) = L(u_1; F_1)L(u_2; F_2),$$

which is just the product of the marginal LCs.

4. In the case of a product distribution, the two-attribute Gini defined in (4) can be written as,

$$1 - G(F_{12}) = [1 - G(F_1)][1 - G(F_2)].$$

The following definition of Lorenz surface was initially proposed by Koshevoy [25] and further results were derived by Koshevoy and Mosler [17,26] and Mosler [27]. Denote by \mathcal{L}_+^2 de set of all 2-dimensional non-negative random vectors X with finite positive marginal expectations. Let $\Psi^{(2)}$ denote the class of all measurable functions from \mathbb{R}_+^2 to $[0, 1]$.

Definition 4. Let $\mathbf{X} \in \mathcal{L}_+^2$. The Lorenz zonoid $L(\mathbf{X})$ of the random vector $\mathbf{X} = (X_1, X_2)^\top$ with distribution F_{12} is defined as,

$$\begin{aligned} L(\mathbf{X}) &= \left\{ \left(\int \psi(x) dF_{12}(x), \int \frac{x_1 \psi(x)}{E(X_1)} dF_{12}(x), \int \frac{x_2 \psi(x)}{E(X_2)} dF_{12}(x) \right) : \psi \in \Psi^{(2)} \right\}, \\ &= \left\{ \left(E[\psi(X)], \frac{E[X_1 \psi(X)]}{E(X_1)}, \frac{E[X_2 \psi(X)]}{E(X_2)} \right) : \psi \in \Psi^{(2)} \right\}. \end{aligned} \tag{5}$$

The Lorenz zonoid is a convex American football-subset of the 3-dimensional unit cube that includes the points $(0,0,0)$ and $(1,1,1)$. Extension to higher dimensions is quite direct. If we consider some bivariate parametric families of distributions, for example the bivariate lognormal distribution (see [28] Section 4.8) used for modelling the joint distribution of household size and income, the computation of previous Lorenz zonoid is not simple. The computation of the formulas involved in $L(\mathbf{X})$ are not explicit in many situations.

2.3. Bivariate Lorenz Surface Based on Bivariate Beta-Generated Distributions

The Arnold’s Lorenz surface can be evaluated implicitly in some relevant bivariate families of income distributions. We present a broad class of Lorenz surfaces based on the class of bivariate beta-generated distributions introduced in [29]. The class of bivariate beta-generated distributions is defined in terms of the joint density by,

$$g_{\mathbf{F}}(x_1, x_2) = k \frac{F_1(x_1)^{a-1} F_2(x_2)^{b-1} (1 - F_1(x_1))^{b+c-1} (1 - F_2(x_2))^{a+c-1} f_1(x_1) f_2(x_2)}{(1 - F_1(x_1) F_2(x_2))^{a+b+c}}, \tag{6}$$

where $k = \frac{\Gamma(a+b+c)}{\Gamma(a)\Gamma(b)\Gamma(c)}$, $F_i(x_i)$ and $f_i(x_i)$, $i = 1, 2$, are genuiene cdfs and pdfs, respectively, and $a, b, c > 0$. The marginal distributions in (6) belong to the class of beta-generated distributions [30] and are given by $X_1 \sim BG(a, c; F_1)$ and $X_2 \sim BG(b, c; F_2)$, which belong to the class \mathcal{L} , that is, the support of F_i , $i = 1, 2$ is $[0, \infty)$ and the first moment exits for both marginals. According to Arnold and Sarabia [31], we define

$$G_{X_1, X_2}^{(1,1)}(x_1, x_2; a, b, c) = \frac{1}{E(X_1 X_2)} \int_0^{x_1} \int_0^{x_2} t_1 t_2 g_{\mathbf{F}}(t_1, t_2) dt_1 dt_2 \tag{7}$$

where $g_{\mathbf{F}}$ is defined in (6), provided that $E(X_1 X_2) < \infty$. Then using Equation 3.5 in Arnold and Sarabia [31] we get the bivariate Lorenz surface,

$$L(u_1, u_2; a, b, c) = G_{X_1, X_2}^{(1,1)}(G_{X_1}^{-1}(x_1), G_{X_2}^{-1}(x_2); a, b, c) \tag{8}$$

Expression (8) includes and generalized several relevant models. For example, if we take $F_i(x_i) = 1 - 1/(1 + x_i)$, $i = 1, 2$ in (6) we obtain the class of bivariate beta distributions of the second kind defined by

$$f_{X_1, X_2}(x_1, x_2; \underline{\alpha}) = \frac{\Gamma(\alpha_1 + \alpha_2 + \alpha_3)}{\Gamma(\alpha_1)\Gamma(\alpha_2)\Gamma(\alpha_3)} \frac{x_1^{\alpha_1-1} x_2^{\alpha_2-1}}{(1 + x_1 + x_2)^{\alpha_1+\alpha_2+\alpha_3}},$$

where $x_1, x_2 > 0$, $\underline{\alpha} = (\alpha_1, \alpha_2, \alpha_3)$ with $\alpha_i > 0$, $i = 1, 2, 3$, which was considered by Arnold and Sarabia [31]. The Lorenz surface is given by,

$$L(u_1, u_2; \alpha_1, \alpha_2, \alpha_3) = F_{X_1, X_2}(F_{X_1}^{-1}(u_1), F_{X_2}^{-1}(u_2); \tilde{\alpha})$$

where $\tilde{\alpha} = (\alpha_1 + 1, \alpha_2 + 1, \alpha_3 - 2)$, with $\alpha_3 > 2$ and $F_{X_1, X_2}(x_1, x_2)$ is the joint cdf of the bivariate beta distributions of the second kind distribution, which can be found in [32].

2.4. Bivariate Generalized Gini Index

In this section we introduce a bivariate version of the generalized Gini index defined by,

$$G^{(\nu)}(X) = 1 - \nu(\nu - 1) \int_0^1 (1 - u)^{\nu-2} L_X(u) du, \tag{9}$$

where $\nu > 1$. If $\nu = 2$, we obtain the usual Gini index. When ν increases, higher weights are attached to small incomes. The limiting case, when ν goes to infinity, depends solely on the lowest income. This definition would be appropriate if we accepted the judgement, introduced by Rawls, that social welfare should be evaluated from the needs the poorest society member. This index was proposed by Donaldson and Weymark [33], Kakwani [34] and was studied by Yitzhaki [35].

Let $(X_1, X_2)^\top$ a bivariate random variable with $E(X_i) < \infty$, $i = 1, 2$, $E(X_1 X_2) < \infty$ and bivariate Lorenz surface $L_{12}(u_1, u_2)$. The bivariate generalized Gini index is defined by,

$$G_{12}^{(\nu_1, \nu_2)}(X_1, X_2) = 1 - k_{\nu_1} k_{\nu_2} \int_0^1 \int_0^1 (1 - u_1)^{\nu_1-2} (1 - u_2)^{\nu_2-2} L_{12}(u_1, u_2) du_1 du_2, \tag{10}$$

where $k_\nu = \nu(\nu - 1)$ and $\nu_1, \nu_2 > 1$. We write $\bar{G}_{12}^{(\nu_1, \nu_2)}(X_1, X_2) = 1 - G_{12}^{(\nu_1, \nu_2)}(X_1, X_2)$. For the special case $\nu_1 = \nu_2 = 2$, we obtain the bivariate Gini index defined in (4). If X_1 and X_2 are independent random variables we have,

$$\bar{G}_{12}^{(\nu_1, \nu_2)}(X_1, X_2) = \bar{G}_1^{(\nu_1)}(X_1) \bar{G}_2^{(\nu_2)}(X_2).$$

3. The Bivariate Sarmanov–Lee Lorenz Surface

The bivariate Lorenz surface (3) admits the following simple explicit representation,

$$L(u_1, u_2; F_{12}) = \frac{1}{E(X_1 X_2)} \int_0^{u_1} \int_0^{u_2} A(x_1, x_2) dx_1 dx_2, \quad 0 \leq u_1, u_2 \leq 1, \tag{11}$$

where

$$A(x_1, x_2) = \frac{F_1^{-1}(x_1) F_2^{-1}(x_2) f_{12}(F_1^{-1}(x_1), F_2^{-1}(x_2))}{f_1(F_1^{-1}(x_1)) f_2(F_2^{-1}(x_2))}. \tag{12}$$

Now, we continue with the so-called bivariate Sarmanov–Lee Lorenz surface and as a previous step, we present the bivariate Sarmanov–Lee distribution.

3.1. The Bivariate Sarmanov–Lee Distribution

Let $\mathbf{X} = (X_1, X_2)^\top$ be a bivariate Sarmanov–Lee (SL) distribution with joint pdf,

$$f(x_1, x_2) = f_1(x_1)f_2(x_2) \{1 + w\phi_1(x_1)\phi_2(x_2)\}, \tag{13}$$

where $f_1(x)$ and $f_2(y)$ are the univariate pdf marginals, $\phi_i(t)$, $i = 1, 2$ are bounded nonconstant functions such that,

$$\int_{-\infty}^{\infty} \phi_i(t)f_i(t)dt = 0, \quad i = 1, 2,$$

and w is a real number which satisfies the condition $1 + w\phi_1(x_1)\phi_2(x_2) \geq 0$ for all x_1 and x_2 . We denote $\mu_i = E(X_i) = \int_{-\infty}^{\infty} tf_i(t)dt$, $i = 1, 2$, $\sigma_i^2 = var(X_i) = \int_{-\infty}^{\infty} (t - \mu_i)^2 f_i(t)dt$, $i = 1, 2$ and $\nu_i = E[X_i\phi_i(X_i)] = \int_{-\infty}^{\infty} t\phi_i(t)f_i(t)dt$, $i = 1, 2$. Properties of this family has been explored by Lee [19]. Moments and regressions of this family can be easily obtained. The product moment is $E(X_1X_2) = \mu_1\mu_2 + w\nu_1\nu_2$, and the regression of X_2 on X_1 is given by

$$E(X_2|X_1 = x_1) = \mu_2 + w\nu_2\phi_1(x_1).$$

The copula associated to (13) is,

$$C(u_1, u_2; w, \phi) = u_1u_2 + \int_0^{u_1} \int_0^{u_2} \tilde{\phi}_1(s_1)\tilde{\phi}_2(s_2)ds_1ds_2,$$

where $\tilde{\phi}_i(s_i) = \phi_i(F_i^{-1}(s_i))$, $i = 1, 2$, and $F_i(x_i)$ are the cdf of \mathbf{X} . The pdf of the copula associated to (13) is,

$$c(u_1, u_2; w, \phi) = \frac{\partial C(u_1, u_2; w, \phi)}{\partial u_1 \partial u_2} = 1 + w\phi_1(F_1^{-1}(u_1))\phi_2(F_2^{-1}(u_2)).$$

Note that (13) and its associated copula has two components: a first component corresponding to the marginal distributions and the second component defining the structure of dependence, given by the parameter w and the functions $\phi_i(u)$, $i = 1, 2$. These two components will be translated to the structure of the associated bivariate Lorenz surface, and the corresponding bivariate Gini index.

In comparison with other copulas, the Sarmanov–Lee copula has several advantages. Its joint pdf and cdf are quite simple and the covariance structure in general is not limited. On the other hand, many of its different probabilistic features can be obtained in a explicit form, which is very useful from a computational point of view. Finally, the SL distribution includes several relevant special cases including the classical Farlie–Gumbel–Morgenstern (FGM) distribution, and the variations proposed in [36,37].

3.2. The Bivariate SL Lorenz Surface

The bivariate SL Lorenz surface [18] is obtained using (13) in Equation (3).

Theorem 1. Let $\mathbf{X} = (X_1, X_2)^\top$ a bivariate Sarmanov–Lee distribution with joint pdf (13) with non-negative marginals satisfying $E[X_1] < \infty$, $E[X_2] < \infty$ and $E[X_1X_2] < \infty$. Then, the bivariate Lorenz surface is given by,

$$L_{SL}(u_1, u_2; F_{12}) = \pi L(u_1; F_1)L(u_2; F_2) + (1 - \pi)L_{g_1}(u_1; F_1)L_{g_2}(u_2; F_2), \tag{14}$$

where

$$\pi = \frac{\mu_1\mu_2}{E(X_1X_2)} = \frac{\mu_1\mu_2}{\mu_1\mu_2 + w\nu_1\nu_2},$$

and $L(u_i; F_i)$, $i = 1, 2$ are the LCs of the marginal distribution X_i , $i = 1, 2$ respectively, and $L_{g_i}(u_i; F_i)$, $i = 1, 2$ represent the concentration curves of the random variables $g_i(X_i) = X_i\phi_i(X_i)$, $i = 1, 2$, respectively.

The interpretation of (14) is the following: the bivariate Lorenz surface can be written as a convex linear combination of two components: a first component corresponding to the product of the marginal LCs, that is the marginal components, and a second component corresponding to the product of the concentration LCs, which corresponds with the structure dependence component.

3.3. Bivariate Generalized Gini Index

The following result provides a convenient write of the two-attribute bivariate generalized Gini defined in (4). This expression generalized previous result by Sarabia and Jordá [18] and permits a simple decomposition of the equality in two factors: a first factor which represent the equality within variables and a second factor corresponding to the equality between variables.

Theorem 2. Let $(X_1, X_2)^\top$ have a bivariate Sarmanov–Lee distribution with bivariate Lorenz surface $L_{12}(u_1, u_2)$. The two-attribute bivariate generalized Gini index defined in (10) admits the following representation,

$$\overline{G}_{12}^{(v_1, v_2)}(X_1, X_2) = \pi \overline{G}_1^{(v_1)}(X_1) \overline{G}_2^{(v_2)}(X_2) + (1 - \pi) \overline{G}_{g_1}^{(v_1)}(X_1) \overline{G}_{g_2}^{(v_2)}(X_2), \tag{15}$$

where $G_i^{(v_i)}(X_i)$, $i = 1, 2$ are the generalized Gini indices of the marginal LCs, and $G_{g_i}^{(v_i)}(X_i)$, $i = 1, 2$ represent the concentration indices of the concentration if the curves $L(u_i, g_i)$, $i = 1, 2$ in the sense of Definition 2.

Proof. The proof is direct using expression (14) and taking into account the definition of bivariate generalized Gini index given in (10). □

Then the overall equality $OE(v_1, v_2)$ can be decomposed in two factors,

$$OE(v_1, v_2) = EW(v_1, v_2) + EB(v_1, v_2), \tag{16}$$

where

$$\begin{aligned} OE(v_1, v_2) &= \overline{G}_{12}^{(v_1, v_2)}(X_1, X_2), \\ EW(v_1, v_2) &= \pi \overline{G}_1^{(v_1)}(X_1) \overline{G}_2^{(v_2)}(X_2), \\ EB(v_1, v_2) &= (1 - \pi) \overline{G}_{g_1}^{(v_1)}(X_1) \overline{G}_{g_2}^{(v_2)}(X_2), \end{aligned}$$

and $EW(v_1, v_2)$ represents the equality within variables and the second factor $EB(v_1, v_2)$ represent the equality between variables. Note that the decomposition the structure of dependence of the underlying bivariate income distribution through the functions g_i , $i = 1, 2$.

4. Bivariate Lorenz Surface Models

In this section we consider several relevant models based on the previous methodology.

4.1. Bivariate Power Lorenz Surfaces Based on the Fgm Family

Let $\mathbf{X} = (X_1, X_2)^\top$ be a bivariate FGM distribution with Power marginals and joint pdf,

$$f(x_1, x_2; a, w) = a_1 a_2 x_1^{a_1 - 1} x_2^{a_2 - 1} \{1 + w(1 - 2x_1^{a_1})(1 - 2x_2^{a_2})\}, \quad 0 \leq x_1, x_2 \leq 1, \tag{17}$$

with $a_i > 0$, $i = 1, 2$ and $-1 \leq w \leq 1$. Note that (17) belongs to the Sarmanov–Lee family with $\phi_i(x_i) = 1 - 2x_i^{a_i}$, $i = 1, 2$ and marginal distributions $F_i(x_i) = x_i^{a_i}$, $i = 1, 2$. The product moment is,

$$E(X_1 X_2) = \frac{a_1 a_2 [(1 + 2a_1)(1 + 2a_2) + w]}{(1 + a_1)(1 + a_2)(1 + 2a_1)(1 + 2a_2)},$$

and the marginal LC, $L(u_i; F_i) = u_i^{1/a_i+1}$, $0 \leq u_i \leq 1$, $i = 1, 2$. After computing the concentration curves and using (14) we obtain,

$$L(u_1, u_2; a, w) = \frac{(1 + 2a_1)(1 + 2a_2)}{(1 + 2a_1)(1 + 2a_2) + w} u_1^{1/a_1+1} u_2^{1/a_2+1} H(u_1, u_2), \tag{18}$$

where

$$H(u_1, u_2) = \left\{ 1 + w \frac{[1 + 2a_1(u_1 - 1) + 2u_1][1 + 2a_2(u_2 - 1) + 2u_2]}{(1 + 2a_1)(1 + 2a_2)} \right\},$$

with $0 \leq u_1, u_2 \leq 1$. The function $H(u, v)$ represents the dependence structure of the bivariate Lorenz surface. Note that (18) can also be written of the form,

$$L(u_1, u_2; a, w) = \frac{u_1^{1/a_1+1} u_2^{1/a_2+1} (A + Bu_1 + Cu_2 + Du_1u_2)}{(1 + 2a_1)(1 + 2a_2) + w}, \quad 0 \leq u_1, u_2 \leq 1,$$

where A, B, C, D are constants which depend on a_1, a_2 and w . Using Equation (4), the bivariate Gini index is given by

$$G(a_1, a_2, w) = \frac{(1 + 3a_1)(1 + 3a_2)(1 + 2a_1 + 2a_2) + wh(a_1, a_2)}{(1 + 3a_1)(1 + 3a_2)[(1 + 2a_1 + 2a_2) + w]},$$

where

$$h(a_1, a_2) = 1 + 3(a_1 + a_2) + [13 - 4(a_1 + a_2) + 4a_1a_2]a_1a_2.$$

If we set $a_1 = a_2 = 1$ we get the bivariate Lorenz surface [38],

$$L_{12}(u_1, u_2, w) = \frac{u_1^2 u_2^2 [9 + w(4u_1 - 3)(4u_2 - 3)]}{9 + w}, \quad 0 \leq u_1, u_2 \leq 1.$$

For this last surface, the bivariate generalized Gini index (10) is given by,

$$\overline{G}^{(v_1, v_2)}(X_1, X_2) = k_{v_1} k_{v_2} \frac{36[w(v_1 - 1)(v_2 - 1) + (3 + v_1)(3 + v_2)]\Gamma(v_1)\Gamma(v_2)}{(9 + w)\Gamma(4 + v_1)\Gamma(4 + v_2)},$$

where $k_\nu = \nu(\nu - 1)$, with $\nu > 1$.

4.2. Bivariate Sarmanov–Lee Lorenz Surfaces with Beta and Gb1 Marginals

Let $X_i \sim \mathcal{B}e(a_i, b_i)$, $i = 1, 2$ be two classical beta distributions with pdf,

$$f_i(x_i; a_i, b_i) = \frac{x_i^{a_i-1} (1 - x_i)^{b_i-1}}{B(a_i, b_i)}, \quad 0 \leq x_i \leq 1, \quad i = 1, 2$$

where $B(a_i, b_i) = \Gamma(a_i)\Gamma(b_i)/\Gamma(a_i + b_i)$, for $i = 1, 2$. This distribution has been proposed as a model of income distribution by McDonald [39], and is commonly used in the literature. If we take the mixing functions $\phi_i(x_i) = x_i - \mu_i$, where $\mu_i = E(X_i) = a_i/(a_i + b_i)$, $i = 1, 2$, the bivariate SL distribution is,

$$f_{12}(x_1, x_2) = f_1(x_1; a_1, b_1) f_2(x_2; a_2, b_2) \left\{ 1 + w \left(x_1 - \frac{a_1}{a_1 + b_1} \right) \left(x_2 - \frac{a_2}{a_2 + b_2} \right) \right\}, \tag{19}$$

where w satisfies,

$$\frac{-(a_1 + b_1)(a_2 + b_2)}{\max\{a_1a_2, b_1b_2\}} \leq w \leq \frac{(a_1 + b_1)(a_2 + b_2)}{\max\{a_1b_2, a_2b_1\}}.$$

A relevant property of this family is that it can be expressed as a linear combination of products of univariate beta densities.

The univariate LC of the univariate classical beta distribution is given by [40],

$$L(u_i; F_i) = G_{\mathcal{B}e(a_i+1, b_i)}[G_{\mathcal{B}e(a_i, b_i)}^{-1}(u_i)], \quad i = 1, 2, \tag{20}$$

where $G_{\mathcal{B}e(a, b)}(z) = (1/B(a, b)) \int_0^z t^{a-1}(1-t)^{b-1} dt$ represents the CDF of a classical Beta distribution. In consequence, the concentration curve can be written as,

$$L_{g_i}(u_i; F_i) = \frac{E(X_i^2)G_{\mathcal{B}e(a_i+2, b_i)}(G_{\mathcal{B}e(a_i, b_i)}^{-1}(u_i)) - E(X_i)^2L(u_i; F_i)}{\text{var}(X_i)}, \quad i = 1, 2. \tag{21}$$

Note that $v_i = E_{F_i}[X_i\phi_i(X_i)] = \text{var}(X_i), i = 1, 2$. Finally, combining (20) with (21) in (14), we obtain the bivariate beta Lorenz surface.

This model can be extended easily to the SL distribution with generalized beta of the first type (GB1) marginals, with pdf,

$$f(x_i; a_i, b_i, p_i) = \frac{p_i x_i^{a_i p_i - 1} (1 - x_i^{p_i})^{b_i - 1}}{B(a_i, b_i)}, \quad 0 \leq x_i \leq 1, \quad i = 1, 2,$$

and mixing function,

$$\phi_i(x_i) = x_i - \mu_i = x_i - \frac{\Gamma(a_i + 1/p_i)\Gamma(a_i + b_i)}{\Gamma(a_i + b_i + 1/p_i)\Gamma(a_i)}, \quad i = 1, 2.$$

4.3. Bivariate SL Lorenz Surfaces with Gamma Marginals

Let $X_i \sim \mathcal{G}(\alpha_i, \lambda_i), i = 1, 2$ be classical gamma distributions with pdf $f_i(x_i) = \lambda_i^{\alpha_i} x_i^{\alpha_i - 1} e^{-\lambda_i x_i} / \Gamma(\alpha_i)$, where $x_i > 0$ and $\alpha_i, \lambda_i > 0, i = 1, 2$. If we consider the mixing function $\phi_i(x_i) = e^{-x_i} - L(1; \alpha_i, \lambda_i)$, where $L(s, \alpha, \lambda) = (1 + s/\lambda)^{-\alpha}$ is the Laplace transform of a gamma distributions [19], we can construct the SL distribution with gamma marginals defined as,

$$f(x_1, x_2) = f_1(x_1)f_2(x_2) \{1 + w (e^{-x_1} - L(1; \alpha_1, \lambda_1)) (e^{-x_2} - L(1; \alpha_2, \lambda_2))\}, \tag{22}$$

with $x_1, x_2 > 0$. Note that (22) can be written as a linear combination of products of univariate gamma distributions. Then, if we define,

$$g_i(x_i) = x_i e^{-x_i} - x_i L(1; \alpha_i, \lambda_i), \quad i = 1, 2,$$

the bivariate SL Lorenz surfaces with gamma marginals is,

$$L_G(u_1, u_2; F_{12}) = \pi L(u_1; F_1)L(u_2; F_2) + (1 - \pi)L_{g_1}(u_1; F_1)L_{g_2}(u_2; F_2),$$

where

$$L(u_i; F_i) = H_{\alpha_i+1}[H_{\alpha_i}^{-1}(u_i)], \quad i = 1, 2,$$

and

$$L_{g_i}(u_i; F_i) = \frac{(1 + \lambda_i)^{-1} H_{\alpha_i+1}[(1 + \frac{1}{\lambda_i})H_{\alpha_i}^{-1}(u_i)] - \lambda_i^{-1} H_{\alpha_i+1}[H_{\alpha_i}^{-1}(u_i)]}{(1 + \lambda_i)^{-1} - \lambda_i^{-1}}, \quad i = 1, 2,$$

being $H_\alpha(x) = \int_0^x t^{\alpha-1} e^{-t} dt / \Gamma(\alpha)$ the cdf of a classical gamma distribution.

4.4. Bivariate SL Lorenz Surfaces with Lognormal Marginals

The lognormal distribution is one of the most relevant distributions in the analysis of income and wealth data. There are several classes of multivariate distributions with lognormal marginals.

The classical multivariate lognormal distribution is defined in terms of a monotone marginal transformation of the classical multivariate normal distribution. Other possible alternatives have been proposed by Sarabia et al. [41] based on conditional specification. In both previous models, we can obtain a bivariate Lorenz surface. If we consider $X_i \sim \mathcal{LN}(\mu_{X_i}, \sigma_{X_i}^2), i = 1, 2$ two lognormal distributions, we define a SL distribution, where the mixing functions are given by,

$$\begin{aligned} \phi_i(x_i) &= f_i(x_i) - \int_0^\infty f_i^2(x_i) dx_i \\ &= f_i(x_i) - \frac{1}{2\sigma_{X_i}\sqrt{\pi}} \exp\left(-\mu_{X_i} + \frac{\sigma_{X_i}^2}{4}\right), \quad i = 1, 2. \end{aligned}$$

If we define $g_i(x_i) = x_i\phi_i(x_i) = x_i f_i(x_i) - x_i \int_0^\infty f_i^2(x_i) dx_i, i = 1, 2$ we can use (14) for obtaining the bivariate Lorenz surface.

4.5. Other Classes of Bivariate Lorenz Surfaces

The class of Lorenz surfaces based on the classical FGM bivariate distribution with Pareto marginals has been considered in [18]. Other alternative families of bivariate Lorenz surfaces can be constructed, including models with marginals specified in terms of univariate LCs [42] and models based on mixture of distributions [43], which permits to incorporate heterogeneity factors in the inequality analysis.

Income variables typically behave differently at the tails of the distribution [44]. In particular, while the generalised beta distribution of the second kind seems to be a suitable model for the bulk of the income distribution [45], incomes above the 95th percentile seem to be better represented by a Pareto distribution [46]. In this regard, the models proposed by Arnold and Sarabia [38], based on mixtures of Lorenz surfaces and bivariate mixture distributions, might be a suitable alternative to model multivariate income distributions.

5. Extensions to Higher Dimensions and Stochastic Dominance

In this section we discuss briefly how to extend the results in previous sections to dimensions higher than two. As well, we include some comments about dominance in higher dimensions.

First, we consider the general definition of Lorenz surface (3). Let $\mathbf{X} = (X_1, \dots, X_m)^\top$ be a random vector in \mathcal{L}_+^m with joint CDF $F_{12\dots m}(x_1, \dots, x_m)$. The multivariate Arnold’s Lorenz surface can be defined as,

$$L(\mathbf{u}; F_{12\dots m}) = \frac{\int_0^{s_1} \dots \int_0^{s_m} \prod_{i=1}^m x_i dF_{12\dots m}(x_1, \dots, x_m)}{E(\prod_{i=1}^m X_i)},$$

where $u_i = \int_0^{s_i} dF_i(x_i), i = 1, 2, \dots, m$ and $0 \leq u_1, \dots, u_m \leq 1$.

The m -dimensional version of the Sarmanov–Lee distribution is defined as,

$$f(x_1, \dots, x_m) = \left\{ \prod_{i=1}^m f_i(x_i) \right\} \left\{ 1 + R_{\phi_1 \dots \phi_m \Omega_m}(x_1, \dots, x_m) \right\}, \tag{23}$$

where

$$\begin{aligned}
 &R_{\phi_1 \dots \phi_m \Omega_m}(x_1, \dots, x_m) \\
 &= \sum_{1 \leq i_1 \leq i_2 \leq m} w_{i_1 i_2} \phi_{i_1}(x_{i_1}) \phi_{i_2}(x_{i_2}) \\
 &= \sum_{1 \leq i_1 \leq i_2 \leq i_3 \leq m} w_{i_1 i_2 i_3} \phi_{i_1}(x_{i_1}) \phi_{i_2}(x_{i_2}) \phi_{i_3}(x_{i_3}) \\
 &+ \dots \\
 &+ w_{12 \dots m} \prod_{i=1}^m \phi_i(x_i),
 \end{aligned}$$

and $\phi_i(x_i), i = 1, \dots, k$ are bounded nonconstant functions such as $\int_{-\infty}^{\infty} \phi_i(t) f_i(t) dt = 0, i = 1, \dots, k$, $\Omega_m = \{w_{i_1 i_2}, w_{i_1 i_2 i_3}, \dots, w_{12 \dots m}\}, w_{i_1 i_2}, w_{i_1 i_2 i_3}, \dots$ are real numbers and $k \geq 2$ and $m \geq 3$. The set of real numbers Ω_m is such that $1 + R_{\phi_1 \dots \phi_m \Omega_m}(x_1, \dots, x_m) \geq 0, \forall (x_1, \dots, x_m) \in \mathbb{R}^m$.

Using expression (23), the m -variate Lorenz surface takes de form,

$$\begin{aligned}
 L(\mathbf{u}; F_{12 \dots m}) &= w_0 \prod_{i=1}^m L(u_i; F_i) + \sum_{i_1, \dots, i_k} w_{i_1 \dots i_k} \prod_{j=i_1}^{i_k} L_{g_j}(u_j; F_j) \prod_{j=i_k+1}^{i_m} L(u_j; F_j) \\
 &+ w_{12 \dots m} \prod_{i=1}^m L_{g_i}(u_i; F_i),
 \end{aligned}$$

where $g_i(x_i) = x_i \phi_i(x_i), i = 1, 2, \dots, m$.

We conclude this section with some references on multidimensional inequality based on zonoid inclusions. Koshevoy [25] and Koshevoy and Mosler [17,26] have studied the partial order of multivariate distributions of transferable attributes induced by the inclusion of either lift zonoids or Lorenz zonoids, defined in (5). In this sense, Koshevoy and Mosler [26] provide the normative rationale for these orderings, using the concept of price majorization (PM) originally proposed by Kolm [9]. See also Andreoli and Zoli [47,48].

Making use of PM order, the distribution of attributes X_1 is always preferred to X_2 by an inequality averse policy maker that is concerned in maximizing the welfare of a distribution of budgets obtained by aggregating the attributes with a given set of prices (see [48]). Mosler [27] provides an interesting result by considering the normative foundations using the extended lift zonoid inclusion order as a robust multidimensional inequality criterion.

From a normative point of view, the interest of the concept of positive PM is related to stochastic orderings that are monotone in attributes and sensitive to the effects of correlation increasing transformations (see Tsui [12,13]). Andreoli and Zoli [48] conceive PM with non-negatives prices as an extension of some multivariate stochastic dominance definition.

6. Application: Multidimensional Inequality in Well-Being

6.1. Data and Estimation Methods

To illustrate the implementation of multidimensional Lorenz surfaces and their associated Gini indices, we collected data from the Demographic and Health Surveys (DHS) of five sub-Saharan African countries: Angola (2015), Uganda (2011), Ethiopia (2011), Haiti (2012) and Nigeria (2012). The DHS are nationally representative household surveys, regularly conducted in more than 85 low- and middle-income countries since 1984. Standard DHS are the largest surveys, designed to cover the national population with sample sizes of between 5000 and 30,000 households.

The DHS were designed to expand the data collected by the World Fertility Surveys and Contraceptive Prevalence Surveys. Hence, these surveys provide rich data on fertility, family planning, health, and nutrition of women and children in developing nations. Although the *Household* and the

Individual modules in the DHS cover the whole population, the data used in this illustration is retrieved from the Household Members Recode. In this module, the age range of the samples are limited to women aged between 15 and 49, but for males the maximum age is not the same in all surveys. Three of the countries considered in this application (Ethiopia, Haiti and Nigeria) include males aged 15–59. Surveys in Angola and Uganda include only men up to the age of 54. The limitation in the age range might introduce a source of bias in our estimates because older population tend to have lower income and worse health, although potentially higher wealth. We are well aware that different age ranges for males might weaken the comparability of the results across countries. In this case, however, this is not a serious issue given the illustrative purpose of the application.

For this illustration, we assume that the level of well-being of the individual i depends solely on two dimensions: wealth (Y_1) and health (Y_2). These two dimensions are included in the set of most prominent aspects of quality of life suggested by Stiglitz et al. [49]. Wealth is captured through a composite index of asset ownership. Compare with income and consumption, household assets seem to provide a better approximation of the economic status of the household. The accurate measurement of household income is extremely difficult due to underreporting and undercoverage issues. This bias is more prevalent in developing countries with large informal sectors. Expenditures are a more reliable measure of household economic status. However, the assessment of consumption expenditures also presents some drawbacks related to the type of expenditures included and the period of time that should be covered. The DHS consider wealth as an underlying unobserved variable, which can be approximated as a linear combination of 30 household assets and utility services associated with a household’s economic position. The weight assigned to each variable is defined using principal component analysis. A potential limitation of this index is the no consideration of financial wealth, which, as a result, underestimates the dispersion of wealth distributions. See [50] for a thorough review on the limitations of the wealth index and their potential solutions.

Health variables in the DHS fit into two categories: those collected objectively and those based on self-reporting. Standard DHS collect data on several nutrition-related biomarkers since the late 90 s. Among them, the hemoglobin level is available in most surveys. Low hemoglobin levels seem to be associated with more physician visits and low body mass index [51]. The prior literature also suggests that there is a close link between low hemoglobin level, malnutrition, and limited access to medical care [52]. Hence, we believe this indicator can be a reliable proxy for the health condition of the population.

We move now to the estimation of the theoretical bivariate distribution of wealth and health. The first step is to normalise wealth and health variables to be bounded between 0 and 1. To do so, we follow the standard procedure used by the United Nations Development Program (UNDP) in the construction of the Human Development Index. The following formula is used to normalise the wealth index and the hemoglobin level:

$$X_i = \frac{Y_i - \min(Y_i)}{\max(Y_i) - \min(Y_i)}, i = 1, 2.$$

The normalised variables are used to estimate the bidimensional distribution of well-being and the level of inequality. Since the variables have support on $[0, 1]$, the classical beta distribution is a suitable model for the marginal distributions. The dependence between dimensions is modelled using the Sarmanov–Lee distribution, studied in Section 4.2.

Let $\mathbf{X} = (X_1, X_2)^\top$ be a bivariate distribution with joint pdf (19) and let $(x_{11}, x_{21}), \dots, (x_{1n}, x_{2n})$ a sample of n individuals from (19). For the estimation of the parameters (a_1, b_1, a_2, b_2, w) we proceed in two steps:

1. Moments estimation of the marginal distributions. We define,

$$m_i = \frac{1}{n} \sum_{j=1}^n x_{ij}, \quad s_i^2 = \frac{1}{n} \sum_{j=1}^n (x_{ij} - m_i)^2, \quad i = 1, 2$$

so, point estimates of the couples $(a_i, b_i), i = 1, 2$ are,

$$\hat{a}_i = \frac{m_i(m_i - m_i^2 - s_i^2)}{s_i^2}, \tag{24}$$

$$\hat{b}_i = \frac{(1 - m_i)(m_i - m_i^2 - s_i^2)}{s_i^2}. \tag{25}$$

2. Moment estimation of the dependence parameter. The estimate of w is based on the simple relation $\rho = w\sigma_1\sigma_2$. Then, if r denotes the sample linear correlation coefficient, and $s_i, i = 1, 2$ the sample standard deviation of the marginal distributions $X_i, i = 1, 2$, the point estimate of w is,

$$\hat{w} = \frac{r}{s_1 \cdot s_2}. \tag{26}$$

Standard errors of $(\hat{a}_i, \hat{b}_i), i = 1, 2$ and w can be easily computed using parametric bootstrap. The Gini index of the marginals distributions is given by,

$$G(\mathcal{B}e(a_i, b_i)) = \frac{\Gamma(a_i + b_i)\Gamma(a_i + \frac{1}{2})\Gamma(b_i + \frac{1}{2})}{\Gamma(a_i + b_i + \frac{1}{2})\Gamma(a_i + 1)\Gamma(b_i)\sqrt{\pi}}, \quad i = 1, 2. \tag{27}$$

6.2. Results

In this section, we present the results of applying the concepts developed in this paper to wealth and health data in five developing countries. Using the methodology described in the previous section we have estimated the parameters of the Sarmanov–Lee distribution considering the beta distribution for modelling marginal distributions (Table 1). Standard errors have been simulated by Monte Carlo with 1000 replications

Table 1. Parameter estimates for the Sarmanov–Lee distribution with classical beta marginals (Equations (24)–(26)). Bootstrap standard errors in parenthesis.

	Income		Health		Dependence
	\hat{a}_1	\hat{b}_1	\hat{a}_2	\hat{b}_2	w
Angola	1.2660 (0.0188)	2.0952 (0.0438)	17.8318 (1.2273)	29.8269 (4.2834)	3.6167 (0.7054)
Ethiopia	2.8269 (0.0758)	5.2741 (0.335)	0.2338 (0.0147)	0.6067 (0.027)	2.3398 (0.114)
Haiti	1.5794 (0.0284)	3.8706 (0.0855)	0.3814 (0.0231)	1.8932 (0.0832)	2.8026 (0.1965)
Nigeria	0.4532 (0.012)	3.1033 (0.0684)	0.1324 (0.0068)	0.3306 (0.0144)	1.1249 (0.1341)
Uganda	0.9473 (0.0185)	2.2020 (0.0539)	0.1994 (0.0211)	0.6314 (0.0534)	1.0116 (0.2055)

Table 2 presents information about the level of well-being inequality in each of the analysed nations. Looking first at unidimensional Gini indices, our estimates suggest that Ethiopia is the most egalitarian country in terms of wealth, with a Gini index of 0.257. The most unequal nation, Nigeria, presents a Gini index of 0.6094. Moving now to health inequality, Angola is by far the country with the lowest Gini index, with a value of 0.1048. Ethiopia and Haiti present a similar level of inequality, around 0.62. Relatively close to this value, we find Nigeria and Uganda, the countries with the highest level of health inequality.

Turning our attention to the bivariate Gini index, our estimates reveal that Angola presents the most equal distribution of well-being if health and wealth are jointly considered. It is followed by Ethiopia, with a Gini index of 0.6921. Then, we find Haiti, where multivariate inequality raises to

0.7273. Uganda and Nigeria are again the most unequal nations, with Gini indices of 0.7954 and 0.864, respectively.

Table 2. Bivariate Gini index (Equation (15)) and marginal Gini indices (27) and Decomposition of the equality (Equation (16)) for the Sarmanov–Lee distribution with classical beta marginals.

	Inequality			Total	Equality	
	Wealth	Health	Bidimensional Gini		Within Variables	Between Variables
Angola	0.3516	0.1048	0.3943	0.6057	0.5766	0.0291
Ethiopia	0.2570	0.6260	0.6921	0.3079	0.2608	0.0471
Haiti	0.3467	0.6214	0.7273	0.2727	0.2294	0.0434
Nigeria	0.6094	0.6678	0.8640	0.1360	0.1174	0.0186
Uganda	0.4198	0.6680	0.7954	0.2046	0.1799	0.0247

Although Angola presents the lowest bivariate Gini index, this country is not the most equal in terms of wealth. This indicates that some sectors of the population (principally the poorest) might be more equally distributed in other countries. To provide a complete picture of the distribution of well-being, it is therefore essential to examine the bidimensional Lorenz surface. Figure 1 shows the bidimensional Lorenz surfaces for wealth and health in Ethiopia and Angola (panel a) and Uganda and Angola (panel b). The comparison of well-being distributions in Ethiopia and Angola reveals that there is not Lorenz dominance, since the Lorenz surface of Ethiopia lies above the Lorenz surface of Angola in the upper-right corner. This suggests that well-being among the poorest individuals with reasonably good health is more equally distributed in Ethiopia. If we now compare the Lorenz surface of Angola with the Lorenz surface of Uganda, the country with the highest bivariate Gini index, we observe a clear pattern of dominance.

We conclude this illustration with the decomposition of the bivariate Gini index (Equations (15) and (16)). Overall equality (expressed as one minus the bivariate Gini index) is decomposed in terms of equality within each dimension and equality between dimensions. The last two columns of Table 2 show the estimates for these two components for the five countries involved in the analysis. These results suggest that multidimensional well-being equality is mainly determined by the within-dimensional distribution, while equality between dimensions represents a residual proportion.

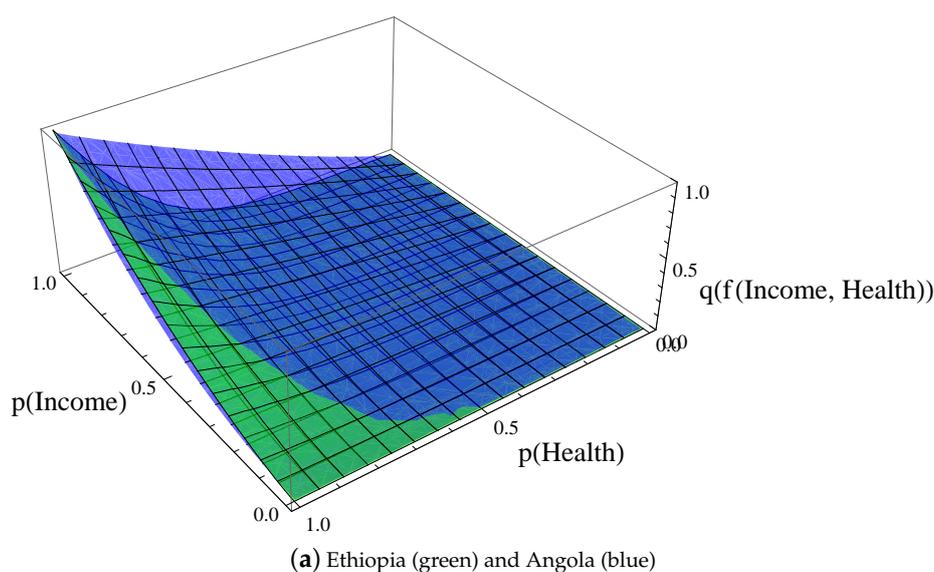


Figure 1. Cont.

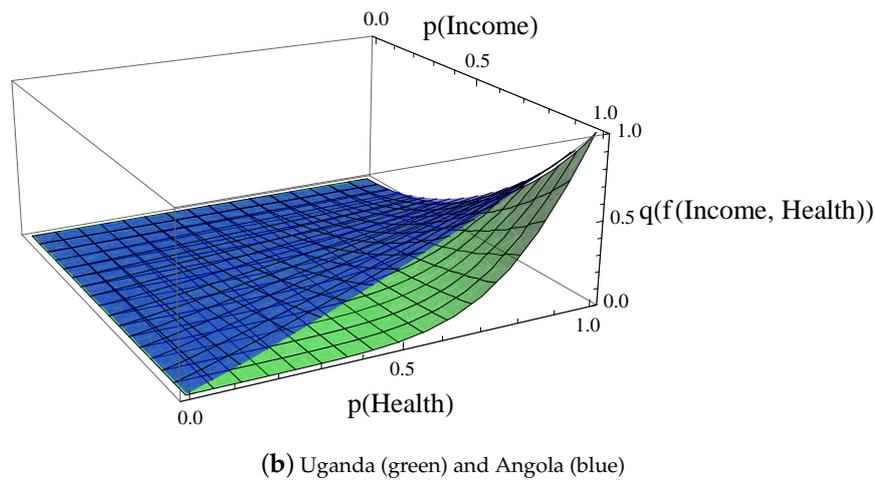


Figure 1. Bivariate Lorenz surfaces: evolution of multidimensional inequality in wealth and health. The curves are defined by Equation (14) with components defined in Equations (20) and (21).

7. Conclusions

In this paper, using the definition proposed by Arnold [16], we presented closed expressions for the bivariate Lorenz surface and its associated Gini index, assuming different formulations of the underlying bivariate income distribution. We discussed a relevant type of models based on the class of bivariate distributions with given marginals described by Sarmanov and Lee [19,20]. For the marginal distributions, We considered several candidates including the Beta, the GB1, the Gamma and the lognormal distributions. Extensions of the bivariate Lorenz surface to higher dimensions were also included. The implementation of multivariate Lorenz surfaces introduced in this paper was illustrated with an application to inequality in well-being in five developing countries.

Author Contributions: Conceptualization, J.M.S.; Data curation, V.J.; Investigation, J.M.S.; Methodology, J.M.S.; Software, and V.J.; Writing—original draft, V.J. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partly funded by the Ministerio de Economía y Competitividad, grant number PID2019-105986GB-C22.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

LC	Lorenz curve
GB1	Generalised beta of the first kind
FGM	Farlie–Gumbel–Morgenstern
SL	Sarmanov–Lee
DHS	Demographic and Health Surveys
UNDP	United Nations Development Program

References

1. Sen, A. The Concept of Development. In *Handbook of Development Economics*; Chenery H., Srinivasan, T.N., Eds.; Elsevier: Amsterdam, The Netherlands, 1988; pp. 9–26.
2. Sen, A. Development as Capabilities Expansion. *J. Dev. Plan.* **1989**, *19*, 41–58.
3. Sen, A. *Development as Freedom*; Oxford University Press: Oxford, UK, 1999.

4. Jorda, V.; López-Noval, B.; Sarabia, J.M. Distributional dynamics of life satisfaction in Europe. *J. Happiness Stud.* **2019**, *20*, 1015–1039. [[CrossRef](#)]
5. Jorda, V.; Sarabia, J.M. International convergence in well-being indicators. *Soc. Indic. Res.* **2015**, *120*, 1–27. [[CrossRef](#)]
6. Atkinson, A.B. Multidimensional deprivation: Contrasting social welfare and counting approaches. *J. Econ. Inequal.* **2003**, *1*, 51–65. [[CrossRef](#)]
7. Atkinson, A.B.; Bourguignon, F. The comparison of multi-dimensioned distributions of economic status. *Rev. Econ. Stud.* **1982**, *49*, 183–201. [[CrossRef](#)]
8. Decancq, K.; Lugo M.A. Inequality of Wellbeing: A Multidimensional Approach. *Economica* **2012**, *79*, 721–746. [[CrossRef](#)]
9. Kolm, S.C. Multidimensional Equalitarianisms. *Q. J. Econ.* **1977**, *91*, 1–13. [[CrossRef](#)]
10. Maasoumi, E. The measurement and decomposition of multi-dimensional inequality. *Econometrica* **1986**, *54*, 991–997. [[CrossRef](#)]
11. Slottje, D.J. Relative Price Changes and Inequality in the Size Distribution of Various Components. *J. Bus. Econ. Stat.* **1987**, *5*, 19–26.
12. Tsui, K.Y. Multidimensional generalizations of the relative and absolute inequality indices: The Atkinson-Kolm-Sen approach. *J. Econ. Theory* **1995**, *67*, 251–265. [[CrossRef](#)]
13. Tsui, K.Y. Multidimensional inequality and multidimensional generalized entropy measures: An axiomatic derivation. *Soc. Choice Welf.* **1999**, *16*, 145–157. [[CrossRef](#)]
14. Taguchi, T. On the two-dimensional concentration surface and extensions of concentration coefficient and Pareto distribution to the two-dimensional case-I. *Ann. Inst. Stat. Math.* **1972**, *24*, 355–382. [[CrossRef](#)]
15. Taguchi, T. On the two-dimensional concentration surface and extensions of concentration coefficient and Pareto distribution to the two-dimensional case-II. *Ann. Inst. Stat. Math.* **1972**, *24*, 599–619. [[CrossRef](#)]
16. Arnold, B.C. *Pareto Distributions*; International Co-Operative Publishing House: Fairland, MD, USA, 1983.
17. Koshevoy, G.; Mosler, K. The Lorenz zonoid of a multivariate distribution. *J. Am. Stat. Assoc.* **1996**, *91*, 873–882. [[CrossRef](#)]
18. Sarabia, J.M.; Jordá, V. Bivariate Lorenz Curves based on the Sarmanov–Lee Distribution. In *Topics in Statistical Simulation, Springer Proceedings in Mathematics & Statistics*; Springer: New York, NY, USA, 2014; Volume 114, pp. 447–455.
19. Lee, M-L.T. Properties of the Sarmanov Family of Bivariate Distributions. *Commun. Stat. Theory Methods* **1996**, *25*, 1207–1222. [[CrossRef](#)]
20. Sarmanov, O.V. Generalized Normal Correlation and Two-Dimensional Frechet Classes. *Doklady* **1966**, *168*, 596–599.
21. Gastwirth, J.L. A general definition of the Lorenz curve. *Econometrica* **1971**, *39*, 1037–1039. [[CrossRef](#)]
22. Kakwani, N.C. Applications of Lorenz Curves in Economic Analysis. *Econometrica* **1977**, *45*, 719–728. [[CrossRef](#)]
23. Taguchi, T. On the structure of multivariate concentration - some relationships among the concentration surface and two variate mean difference and regressions. *Comput. Stat. Data Anal.* **1988**, *6*, 307–334. [[CrossRef](#)]
24. Arnold, B.C. *Majorization and the Lorenz Curve*; Lecture Notes in Statistics 43; Springer: New York, NY, USA, 1987.
25. Koshevoy, G. Multivariate Lorenz majorization. *Soc. Choice Welf.* **1995**, *12*, 93–102. [[CrossRef](#)]
26. Koshevoy, G.; Mosler, K. Multivariate Gini indices. *J. Multivar. Anal.* **1997**, *60*, 252–276. [[CrossRef](#)]
27. Mosler, K. *Multivariate Dispersion, Central Regions and Depth: The Lift Zonoid Approach*; Springer: Berlin, Germany, 2002.
28. Kleiber, C.; Kotz, S. *Statistical Size Distributions in Economics and Actuarial Sciences*; John Wiley: Hoboken, NJ, USA, 2003.
29. Sarabia, J.M.; Prieto, F.; Jordá, V. Bivariate beta-generated distributions with applications to well-being data. *J. Stat. Distrib. Appl.* **2014**, *1*, 1–15. [[CrossRef](#)]
30. Alexander, C.; Cordeiro, G.M.; Ortega, E.M.M.; Sarabia, J.M. Generalized beta-generated distributions. *Comput. Stat. Data Anal.* **2012**, *56*, 1880–1897. [[CrossRef](#)]
31. Arnold, B.C.; Sarabia, J.M. Analytic Expressions for Multivariate Lorenz Surfaces. *Sankhya A Indian J. Stat.* **2018**, *80*, 84–111. [[CrossRef](#)]
32. Balakrishnan, N.; Lai, C.-D. *Continuous Bivariate Distributions*; Springer: New York, NY, USA, 2009.

33. Donaldson, D.; Weymark, J.A. A single parameter generalization of the Gin index of inequality. *J. Econ. Theory* **1980**, *22*, 67–86. [[CrossRef](#)]
34. Kakwani, N.C. *Income Inequality and Poverty, Methods and Estimation and Policy Applications*; Oxford University Press: New York, NY, USA, 1980.
35. Yitzhaki, S. On an extension of the Gini inequality index. *Int. Econ. Rev.* **1983**, *24*, 617–628. [[CrossRef](#)]
36. Huang, J.S.; Kotz, S. Modifications of the Farlie-Gumbel-Morgenstern distributions. A tough hill to climb. *Metrika* **1999**, *49*, 135–145. [[CrossRef](#)]
37. Bairamov, I.; Kotz, S. On a new family of positive quadrant dependent bivariate distributions. *Int. Math. J.* **2003**, *3*, 1247–1254.
38. Arnold, B.C.; Sarabia, J.M. *Majorization and the Lorenz order with Applications in Applied Mathematics and Economics*; Springer: New York, NY, USA, 2018.
39. McDonald, J.B. Some generalized functions for the size distribution of income. *Econometrica* **1984**, *52*, 647–663. [[CrossRef](#)]
40. Sarabia, J.M. Parametric Lorenz Curves: Models and Applications. In *Modeling Income Distributions and Lorenz Curves*; Chotikapanich, D., Ed.; Springer: New York, NY, USA, 2008; pp. 167–190.
41. Sarabia, J.M.; Castillo, E.; Pascual, M.; Sarabia, M. Bivariate Income Distributions with Lognormal Conditionals. *J. Econ. Inequal.* **2007**, *5*, 371–383. [[CrossRef](#)]
42. Sarabia, J.M.; Castillo, E.; Slottje, D. An Ordered Family of Lorenz Curves. *J. Econ.* **1999**, *91*, 43–60. [[CrossRef](#)]
43. Sarabia, J.M.; Castillo, E.; Pascual, M.; Sarabia, M. Mixture Lorenz Curves. *Econ. Lett.* **2005**, *89*, 89–94. [[CrossRef](#)]
44. Hlasny, V. Nonresponse Bias in Inequality Measurement: Cross-Country Analysis Using Luxembourg Income Study Surveys. *Soc. Sci. Q.* **2020**, *101*, 712–731. [[CrossRef](#)]
45. Jorda, V.; Sarabia, J.M.; Jäntti, M. Estimation of income inequality from grouped data. *arXiv* **2018**, arXiv:1808.09831.
46. Jenkins, S.P. Pareto models, top incomes and recent trends in UK income inequality. *Economica* **2017**, *84*, 261–289. [[CrossRef](#)]
47. Andreoli, F.; Zoli, C. *Measuring Dissimilarity*; Working Paper Series; Department of Economics University of Verona: Verona, Italy, 2014; Volume 23.
48. Andreoli, F.; Zoli, C. From unidimensional to multidimensional inequality: A review. *Metron* **2020**, *78*, 5–42. [[CrossRef](#)]
49. Stiglitz, J.E.; Sen, A.; Fitoussi, J.P. *Report by the Commission on the Measurement of Economic Performance and Social Progress*; Citeseer: University Park, PA, USA, 2009.
50. Hlasny, V.; AlAzzawi, S. Asset inequality in the MENA: The missing dimension? *Q. Rev. Econ. Financ.* **2019**, *73*, 44–55. [[CrossRef](#)]
51. Balarajan, Y.; Ramakrishnan, U.; Ozaltin, E.; Shankar, A.H.; Subramanian, S.V. Anaemia in low-income and middle-income countries. *Lancet* **2011**, *378*, 2123–2135. [[CrossRef](#)]
52. Scanlon, K.S.; Yip, R.; Schieve, L.A.; Cogswell, M.E. High and low hemoglobin levels during pregnancy: Differential risks for preterm birth and small for gestational age. *Obstet. Gynecol.* **2000**, *96*, 741–748. [[CrossRef](#)]

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).