Working Paper No. 305

Sequential Comparisons of Generalized Lorenz Curves for Different Demographics

Kazuyuki Nakamura

March, 2017

UNIVERSITY OF TOYAMA 富山大学

FACULTY OF ECONOMICS
UNIVERSITY OF TOYAMA

Sequential Comparisons of Generalized Lorenz Curves for Different Demographics

Kazuyuki Nakamura
March, 2017

Abstract

Jenkins and Lambert (1993) and Chambaz and Maurin (1998) proposed extensions of the sequential generalized Lorenz dominance (SGL) criterion that was proposed by Atkinson and Bourguignon (1987): the extended version of SGL made it possible to compare distributions for different demographics. However, the tests to check the extended SGL “are not expressible in terms of generalized Lorenz curves” (Lambert 2001, p.79). In this paper, we show that the dominance condition can be easily checked by sequential comparisons of a modified version of the generalized Lorenz curve. We apply this procedure by comparing the income distributions of Italian households using data obtained from the Survey on Household Income and Wealth (SHIW, Bank of Italy) from 2006 and 2012.

Keywords: Sequential generalized Lorenz dominance; welfare ordering; different demographics

JEL codes: D31, D63.

1 Introduction

In the literature of income inequality, the heterogeneity of individuals or households is an important issue. For a population that is classified into various groups that are homogeneous in attributes other than income, Atkinson and Bourguignon (1987) proposed a set of dominance criteria, including the sequential generalized Lorenz dominance (SGL)[1].

The SGL is attractive for several reasons. First, the SGL enables us to obtain welfare orderings without rigorous specification of a utility function. In particular, when heterogeneity among households arises from a difference in household size, the SGL can provide a welfare ordering without relying on specific equivalence scales. Second, the SGL criterion can be easily checked by using the familiar generalized Lorenz (GL) curves that were proposed by Shorrocks

[1] In the subsequent analysis, the characteristics of the SGL ordering have been considered. Ok and Lambert (1999) revealed that a social welfare function (SWF) supports the SGL ordering if and only if it is needs-based. Ebert (2000) provided a characterization of SGL ordering based on transfer principles. Ooghe (2007) argued that the SGL ordering is equivalent to the welfare orderings under the utilitarian SWFs which gives a larger weight to the needy households.

(1983), sequentially.

The SGL is applicable when we wish to predict the distributional impact induced by a policy change.[2] However, the SGL can only be used for comparison when considering a fixed demographic composition. It cannot be used to compare the distributions of groups that have different compositions. In practice, international or intertemporal income distributions are often compared; these usually have different demographic compositions.

Jenkins and Lambert (1993) extended the SGL to cases in which there are demographic differences. They imposed an additional restriction on the utility function and showed a sufficient condition for welfare dominance in the presence of demographic differences. Hereafter we refer to their dominance criterion as the *extended SGL*. Chambaz and Maurin (1998) showed that this sufficient condition is a necessary condition for welfare dominance.

These analyses rely on the second-order stochastic dominance condition in order to compare income distributions.[3] Thus, the analytical procedure that checks the extended SGL uses distribution functions rather than the GL curves. Indeed, in his informative textbook, Lambert (2001) argued that the tests to confirm the extended SGL "require the analyst to use numerical integration techniques to compute areas under distribution functions" (Lambert, 2001. p.79).

In this paper, we show that the dominance conditions proposed by Jenkins and Lambert (1993) and Chambaz and Maurin (1998) can be easily checked by using a slightly modified version of the GL curves. That is, as in an SGL comparison with unchanged demographic compositions, the welfare-dominance relation can be investigated without the comparison of distribution functions, even in the presence of demographic differences. In addition to welfare ordering, Jenkins and Lambert (1993) and Chambaz and Maurin (1998) applied their dominance conditions to the analysis of poverty. Our procedure can also be used to check the poverty gap dominance when there are demographic differences.

Our procedure for checking the dominance relation is quite simple. When comparing two populations with unequal numbers of low-income households, we add virtual households to the smaller population so that the two are equalized. Using these adjusted distributions, we compare their welfare by using the GL curves. This procedure is not only simple, but the results can be shown visually by using the familiar GL curve.

As mentioned above, the procedures proposed in this paper add nothing to the existing dominance criteria. Rather, we make clearer the implication of the extended SGL criterion by providing its simple representation. The SGL criterion is said to be too robust to compare

---

[2] For example, López-Laborda and Onrubia (2005) applied the idea of the SGL to a regional context and considered the decentralization of a tax system that results in better regional welfare than did the initial unified tax system.

[3] Based on the first-, second- and third-order stochastic dominance conditions, Lambert and Ramos (2002) provide the sequential tests for checking welfare dominance when there are demographic differences.

income distributions although its generality corresponds to broad class of utility functions[4]. Since the procedure proposed in this paper can be easily implemented, researchers who employ the dominance criteria other than the extended SGL can confirm the robustness of their judgements. On the other hand, Moyes (2012) argued that the welfare implication by the extended SGL highly depends on the maximum level of conceivable income. The graphical representation by the GL curves can clearly show the change in welfare judgements induced by a change in the maximum conceivable income.

The reminder of this paper is organized as follows. Section 2 provides the analytical framework. In Section 3, we will show the results for the welfare dominance relation. In Section 4, we consider its application for poverty analysis. In Section 5, the proposed procedures are applied to the comparison of the distributions of Italian households. In the last section, we present some closing remarks.

2 Analytical Framework

Consider two societies *X* and *Y*, consisting of $n(>1)$ households[5]. In each society, households are classified into $H(2 \leq H \leq n)$ groups according to their needs for income. Let $\mathcal{H} = \{1, \dots, H\}$ be the set of groups, where $h \in \mathcal{H}$ denotes the h-th needy group. Let $\mathcal{N}(h)$ be the set of households who belong to the h-th group. We denote by $N(h) \geq 0$ the number of households classified into the h-th group: $N(h) = \#\mathcal{N}(h)$. Furthermore, denote by $\widetilde{\mathcal{N}}(h) = \bigcup_{j=1}^{h} \mathcal{N}(j)$ the set of households having a type lower or equal than h. It follows that $\widetilde{N}(h) = \#\widetilde{\mathcal{N}}(h)$ is such that $\widetilde{N}(h) = \sum_{j=1}^{h} N(j)$. Obviously, $\widetilde{N}(H) = \sum_{h \in \mathcal{H}} N(h) = n$. Hereafter, without loss of generality, we set $\mathcal{N}(h) = \{\widetilde{N}(h-1)+1, \dots, \widetilde{N}(h)\}$ for $h \in \mathcal{H}$ and $\widetilde{N}(0) = 0$.

The distribution of income in society *X* can be represented by an *n*-dimensional row vector $\mathbf{x} = (x_1, \dots, x_n)$, where $x_i \in [0, \bar{v}]$ is the income of the *i*-th household, and $\bar{v}$ is the maximum conceivable income level. The distribution of income in the h-th group can be represented by an $N_X(h)$-dimensional vector $\mathbf{x}^h = (x_{\widetilde{N}_X(h-1)+1}, \dots, x_{\widetilde{N}_X(h)})$. Combining $\mathbf{x}^h$s, we denote by $\tilde{\mathbf{x}}^h$ income profile of households whose needs are no more than h: $\tilde{\mathbf{x}}^h = \left(x_1, \dots, x_{\widetilde{N}_X(h)}\right)$. Obviously, $\tilde{\mathbf{x}}^H = \mathbf{x}$ holds. For the later discussions, we denote by $\tilde{\mathbf{x}}_\uparrow^h := (x_{(1)}^h, \dots, x_{(\widetilde{N}_X(h))}^h)$ the vector obtained by rearranging the entries of $\tilde{\mathbf{x}}^h$ into ascending order[6]: $x_{(1)}^h \leq, \dots, \leq$

[4] For detail discussion, see Fleurbaey et al. (2003).

[5] While Jenkins and Lambert (1993) and Chambaz and Maurin (1998) focus on the continuous distribution, present analysis is based on discrete distribution of income. As discussed in Fishburn and Vickson (1978) and Muliere and Scarsini (1989), present analysis is compatible with existing results based on continuous distribution.

[6] That is, $\tilde{\mathbf{x}}_\uparrow^h$ is obtained by some permutation matrix Π as $\tilde{\mathbf{x}}_\uparrow^h = \tilde{\mathbf{x}}^h \Pi$. It should be noted that $x_{(j)}^h = x_{(j)}^k$ does not necessarily hold for $h, k \in \mathcal{H}$ and $j \in \widetilde{\mathcal{N}}_X(h) \cap \widetilde{\mathcal{N}}_X(k)$.

$x^h_{(\tilde{N}_X(h))}$. Similarly, the income distribution in society *Y* can be represented by an *n*-dimensional vector $\boldsymbol{y} = (y_1, \ldots, y_n)$ and $y_i \in [0, \bar{v}]$, which is decomposed according to the level of the need, that is, $\mathbf{y} = [\mathbf{y}^1, \ldots, \mathbf{y}^H]$, where $\mathbf{y}^h = (y_{\tilde{N}_Y(h-1)+1}, \ldots, y_{\tilde{N}_Y(h)})$ is the $N_Y(h)$-dimensional vector. As in society *X*, we define vectors $\tilde{\mathbf{y}}^h$ and $\tilde{\mathbf{y}}^h_\uparrow$.

3 Welfare Dominance and GL Comparisons

We denote by $U(y,h)$ the social valuation – utility hereafter – of an household having income $y \in [0,+\infty)$ and $h \in \mathcal{H}$. The household utility function is continuous and twice differentiable with respect to income. We denote by $\mathcal{U}$ the set of such functions. Jenkins and Lambert (1993), and Chambaz and Maurin (1998) assume the following hypotheses, where U_1 is the first derivative of U according to its first argument:

U1. $$U_1(y,1) \geq U_1(y,2) \geq, \ldots, \geq U_1(y,H) \geq 0, \quad \forall y \leq \bar{v},$$

U2. $$U(\bar{v},1) = U(\bar{v},2) =, \ldots, = U(\bar{v},H),$$

U3. $$U_{11}(y,1) \leq U_{11}(y,2) \leq, \ldots, \leq U_{11}(y,H) \leq 0, \quad \forall y \leq \bar{v}.$$

According to Jenkins and Lambert (1993), and Chambaz and Maurin (1998), we introduce the following class of utility function:

$$\mathcal{U}_{JL} = \{U \in \mathcal{U}\colon \text{U1, U2 and U3 are satisfied}\}.$$

We then assume that the social planner is endowed with a utilitarian social welfare function:

$$W_X = \frac{1}{n} \sum_{h \in \mathcal{H}} \sum_{i \in \mathcal{N}_X(h)} U(x_i, h). \tag{1}$$

For society Y, we define by $W_Y = (1/n) \sum_{h \in \mathcal{H}} \sum_{i \in \mathcal{N}_Y(h)} U(y_i, h)$ the social welfare.

Let $F_X(x|h)$ be the conditional cumulative distribution function at income $x \in [0,+\infty)$ and for type $h \in \mathcal{H}$. Chambaz and Maurin (1998) have shown that the following two conditions are equivalent (Chambaz and Maurin, 1998, Proposition 2).

$$W_X \geq W_Y \ \ \forall U \in \mathcal{U}_{JL}, \tag{A}$$

$$\sum_{j=1}^{h} \int_0^t \left[q_{j,X} F_X(\xi|j) - q_{j,Y} F_Y(\xi|j)\right] d\xi \leq 0 \quad \forall t \leq \bar{v}, \forall h \in \mathcal{H}. \tag{B}$$

where $q_{j,X} = N_X(j)/n$ denotes the household share of the type $j \in \mathcal{H}$.

Moyes (2012) gives a simple representation of condition (B) by using joint distribution function. Let $F(x, h)$ be the joint distribution function at income $x \in [0, +\infty)$ and for type $h \in \mathcal{H}$. Noting that $F_X(x, h) = \sum_{j=1}^{h} q_{j,X} F_X(x|j)$, we can replace (B) with the following condition.

$$\int_0^t [F_X(\xi, h) - F_Y(\xi, h)] d\xi \leq 0 \ \ \forall t \leq \bar{v}, \forall h \in \mathcal{H}. \tag{B'}$$

Since our concern is discrete distribution, condition (B') can be rewritten as follows[7]:

$$\frac{1}{n}\left[\sum_{i \in \widetilde{N}_X(h)} (t - x_i)^+ - \sum_{i \in \widetilde{N}_Y(h)} (t - y_i)^+\right] \leq 0 \ \ \forall t \leq \bar{v}, \forall h \in \mathcal{H}, \tag{B2}$$

where $(a)^+ = \max\{0, a\}$. If we consider the situation of $\widetilde{N}_X(h) = \widetilde{N}_Y(h) \ \ \forall h \in \mathcal{H}$, the GL comparison is directly applicable. To connect condition (B) to the GL dominance criterion, we present the following Lemma.

Lemma 1 The following condition is equivalent to condition (B2).

$$\sum_{i \in \widetilde{N}_X(h)} (t - x_i)^+ - \sum_{i \in \widetilde{N}_Y(h)} (t - y_i)^+ - \left(\widetilde{N}_X(h) - \widetilde{N}_Y(h)\right)(t - \bar{v})^+ \leq 0, \ \ \forall t \in [0, \infty), \ \ \forall h \in \mathcal{H}. \tag{B3}$$

Proof First, (B3) is equivalent to (B2) if we consider the case of $t \leq \bar{v}$ in (B3). For $t \in (v, \infty)$, (B3) is reduced to $\sum_{i \in \widetilde{N}_X(h)} (\bar{v} - x_i)^+ - \sum_{i \in \widetilde{N}_Y(h)} (\bar{v} - y_i)^+ \leq 0$, which implies (B2). ■

[7] For example, see Moyes (1999).

Now, in order to connect the result by the Chambaz and Maurin (1998) to the GL curves comparison, we modify the income vectors. Let $\mathbf{v}^h = (\bar{v}, \dots, \bar{v})$ be a $\left|\tilde{N}_X(h) - \tilde{N}_Y(h)\right|$ dimensional vector whose entries are all equal to $\bar{v}$. Let $\hat{N}(h) = \max\{\tilde{N}_X(h), \tilde{N}_Y(h)\}$. By using $\mathbf{v}^h$, we define the following vector for $h \in \mathcal{H}$.

$$\hat{\mathbf{x}}^h = (\hat{x}_1^h, \dots, \hat{x}_{\hat{N}(h)}^h) = \begin{cases} \tilde{\mathbf{x}}_\uparrow^h & if\ \tilde{N}_X(h) \geq \tilde{N}_Y(h) \\ \left[\tilde{\mathbf{x}}_\uparrow^h, \mathbf{v}^h\right] & if\ \tilde{N}_X(h) < \tilde{N}_Y(h) \end{cases} \tag{2a}$$

That is, $\hat{\mathbf{x}}^h$ is a $\hat{N}(h)$-dimensional vector whose j-th entry, $\hat{x}_j^h = \tilde{x}_{(j)}^h$ for $j \in \tilde{\mathcal{N}}_X(h)$. If $\tilde{N}_X(h) < \tilde{N}_Y(h)$, then $\hat{x}_j^h = \bar{v}$ for $j = \tilde{N}_X(h) + 1, \dots, \hat{N}(h)$. Since $\bar{v}$ is possible maximum income, $\hat{x}_1^h \leq, \dots, \leq \hat{x}_{\hat{N}(h)}^h$ holds. Similarly, we define

$$\hat{\mathbf{y}}^h = (\hat{y}_1^h, \dots, \hat{y}_{\hat{N}(h)}^h) = \begin{cases} \tilde{\mathbf{y}}_\uparrow^h & if\ \tilde{N}_X(h) \leq \tilde{N}_Y(h) \\ \left[\tilde{\mathbf{y}}_\uparrow^h, \mathbf{v}^h\right] & if\ \tilde{N}_X(h) > \tilde{N}_Y(h) \end{cases} \tag{2b}$$

The following result is well-known in the theory of majorization[8].

Lemma 2 Condition (B3) is equivalent to the following condition.

$$\sum_{j=1}^{r} \hat{x}_j^h \geq \sum_{j=1}^{r} \hat{y}_j^h, \qquad r = 1, \dots, \hat{N}(h), \forall h \in \mathcal{H}. \tag{B4}$$

Proof See Marshall et al. (2011, Proposition B.4, p.158). ■

Now, we consider the GL curve for the virtual distribution $\hat{\mathbf{x}}^h$. As well-known in the literature, the GL curve is given by the second-order inverse distribution function for a given distribution[9]. Let $GL(\mathbf{x}, p)$ for $p \in [0,1]$ be the GL function for $\mathbf{x}$. According to Moyes (1999), we can represent the GL function as follows:

$$GL\left(\hat{\mathbf{x}}^h, p\right) = \sum_{j=1}^{k(p;\hat{N}(h))} \left(p - \frac{j-1}{\hat{N}(h)}\right)\left(\hat{x}_j^h - \hat{x}_{j-1}^h\right) \text{ for } p \in [0,1] \text{ and } h \in \mathcal{H}, \tag{3}$$

with $\hat{x}_0^h = 0$, where $k \coloneqq \min\{k \in \{1, .2, \dots, \hat{N}(h)\}: p \leq k/\hat{N}(h)\}$. Similarly, we can define the

[8] In the theory of majorization, it is said that $\hat{\mathbf{y}}^h$ weakly super-majorizes $\hat{\mathbf{x}}^h$ if and only if (B4) holds (e.g., Marshall et al. 2011).

[9] For example, see Gastwirth(1971).

GL function for $\hat{\mathbf{y}}^h$ as $GL(\hat{\mathbf{y}}^h, p)$.

Our main result can be summarized by the following proposition.

Proposition 1 The following two conditions are equivalent.

$$W_X \geq W_Y, \forall U \in \mathcal{U}_{JL}, \tag{A}$$

$$GL(\hat{\mathbf{x}}^h, p) \geq GL(\hat{\mathbf{y}}^h, p) \ \ \forall p \in [0,1], \forall h \in \mathcal{H}. \tag{B5}$$

Proof From Lemmas 1 and 2, it is confirmed that (A)⇔(B4). Thus, we will prove (B4)⇔(B5). We can rewrite (3) as follows:

$$GL(\hat{\mathbf{x}}^h, p) = \frac{1}{\widehat{N}(h)} \left\{ \alpha\left(p; \widehat{N}(h)\right) \sum_{j=1}^{k(p;\widehat{N}(h))} \hat{x}_j^h + \left[1 - \alpha\left(p; \widehat{N}(h)\right)\right] \sum_{j=0}^{k(p;\widehat{N}(h))-1} \hat{x}_j^h \right\}, \tag{4}$$

where $\alpha(p; \widehat{N}(h)) \equiv \widehat{N}(h)p - k(p; \widehat{N}(h)) + 1 \in [0, 1]$. From (4), it is obvious that (B5) holds if (B4) is fulfilled. Next, suppose that $\sum_{j=1}^{r^+} \hat{x}_j^{h^+} < \sum_{j=1}^{r^+} \hat{y}_j^{h^+}$ for some $r^+ \in \{1, \dots, \widehat{N}(h^+)\}$ and $h^+ \in \mathcal{H}$. Taking $p = r^+/\widehat{N}(h^+)$, we have $k(r^+/\widehat{N}(h^+); \widehat{N}(h^+)) = r^+$ and $\alpha(r^+/\widehat{N}(h^+); \widehat{N}(h^+)) = 1$. Therefore, $GL(\hat{\mathbf{x}}^{h^+}, r^+/\widehat{N}(h^+)) < GL(\hat{\mathbf{y}}^{h^+}, r^+/\widehat{N}(h^+))$. ■

Proposition 1 implies that the dominance criterion discussed by Chambaz and Maurin (1998) can be easily checked by sequential comparisons of the modified GL curves.

Example 1 Suppose we have the following distributions of income partitioned into three subgroups:

$$\mathbf{x}^1 = (2, 5, 8), \qquad \mathbf{x}^2 = (4, 6, 10, 12), \qquad \mathbf{x}^3 = (3, 7, 9),$$

$$\mathbf{y}^1 = (2, 3, 8, 9), \qquad \mathbf{y}^2 = (1, 4), \qquad \mathbf{y}^3 = (5, 6, 7, 15).$$

Suppose that $\bar{v} = 15$. Let us consider the dominance condition stated in Proposition 1. We can draw the GL curves of $\hat{\mathbf{x}}^h$ and $\hat{\mathbf{y}}^h$ for $h \in \{1,2,3\}$ as in Figure 1 (*a*)–(*c*), where the dotted lines in (*a*) and (*b*) correspond to $\bar{\mathbf{v}}$. It is clear from Figure 1 that the dominance condition is met. Furthermore, since $\sum_{i=1}^{\tilde{n}_X^2} x_i - \sum_{i=1}^{\tilde{n}_Y^2} y_i = 20$, we can confirm that $W_X \geq W_Y$ holds for $\bar{v} \leq 20$.

[Figure 1 placed here]

When investigating the welfare dominance condition, it is important to determine to what extent the differences in needs should be considered. In the above example, if we allow for the difference in utility between the second and third needy groups in income larger than 20, the welfare dominance condition is violated. As an extreme case, if $\bar{v}$ approaches to infinity, $\mathbf{x}$ is never preferred to $\mathbf{y}$ in the sense of the extended SGL unless $\tilde{N}_X^h \leq \tilde{N}_Y^h$ holds for all $h \in \mathcal{H} \backslash \{H\}$. We will discuss this point in the later section.

4 Poverty Dominance

The argument presented above can be applied to the analysis of poverty. First, let $\bar{z}_h \in \mathbb{R}_+$ be the poverty line for the households in type $h \in \mathcal{H}$. According to the literature, we will consider the profiles of the poverty lines in decreasing order.[10] That is, $\bar{z}_1 \geq, \dots, \geq \bar{z}_H$. Let $\mathcal{Z}$ be a set of *H*-dimensional vectors $\mathcal{Z} \equiv \{(\bar{z}_1, \dots, \bar{z}_H): \bar{z}_1 \geq, \dots, \geq \bar{z}_H\}$. The profile of poverty line is represented as $\bar{\mathbf{z}} = (\bar{z}_1, \dots, \bar{z}_H) \in \mathcal{Z}$.

For society *X*, we consider the following poverty measure:

$$P_X(\bar{\mathbf{z}}) = \sum_{h \in \mathcal{H}} \sum_{i \in \mathcal{N}_X(h)} P(x_i, z_h), \qquad \bar{\mathbf{z}} \in \mathcal{Z}, \tag{5}$$

For society *Y*, we consider the poverty measure $P_Y(\bar{\mathbf{z}}) = \sum_{h \in \mathcal{H}} \sum_{i \in \mathcal{N}_Y(h)} P(y_i, z_h)$. The function $P(y, z)$ is continuous and twice differentiable with respect to income. We denote by $\mathcal{P}$ the set of such functions. According to Chambaz and Maurin (1998), we assume the following properties, where P_1 is the first derivative of P with respect to its first argument.

P1. $$P_1(y, z_1) \leq P_1(y, z_2) \leq, \dots, \leq P_1(y, z_H) \leq 0, \ \ \forall y \leq \bar{v},$$

P2. $$\begin{cases} P(y, z_h) \geq 0 & if\ y \leq z_h, \quad \forall y \leq \bar{v}\ and\ \forall h \in \mathcal{H}, \\ P(y, z_h) = 0 & if\ y > z_h, \quad \forall y \leq \bar{v}\ and\ \forall h \in \mathcal{H}, \end{cases}$$

P3. $$P_{11}(y, z_1) \geq P_{11}(y, z_2) \geq, \dots, \geq P_{11}(y, z_H) \geq 0, \ \ \forall y \leq \bar{v}.$$

[10] For example, see Atkinson (1992), Jenkins and Lambert (1993), and Chambaz and Maurin (1998).

All assumptions are employed in Chambaz and Maurin (1998). For a given profile of poverty lines $\bar{\mathbf{z}} \in \mathcal{Z}$, we will denote as $\mathcal{P}_{CM}(\bar{\mathbf{z}})$ the class of the poverty indices that satisfy P1-P3.

$$\mathcal{P}_{CM}(\mathbf{z}) = \{P \in \mathcal{P}: \text{P1, P2 and P3 are satisfied and } \mathbf{z} \in \mathcal{Z}\}.$$

Poverty dominance condition presented by Chambaz and Maurin (1998) states that the following two conditions are equivalent (Chambaz and Maurin, 1998, Proposition 4).

$$P_X \leq P_Y, \forall P \in \mathcal{P}_{CM}(\bar{\mathbf{z}}), \bar{\mathbf{z}} \in \mathcal{Z}, \tag{C}$$

$$\sum_{j=1}^{h} \int_0^t \left[q_{j,X} F_X(\xi|j) - q_{j,Y} F_Y(\xi|j)\right] d\xi \leq 0, \quad \forall t \leq \bar{z}_h, \forall h \in \mathcal{H} \textit{ and } \bar{\mathbf{z}} \in \mathcal{Z}. \tag{D}$$

As in condition (B), noting that $\sum_{j=1}^{h} \int_0^t q_{j,X} F_X(\xi|j) d\xi = \int_0^t F_X(\xi, h) d\xi$ holds, we can rewrite (D) as

$$\frac{1}{n}\left[\sum_{i \in \widetilde{N}_X(h)} (t - x_i)^+ - \sum_{i \in \widetilde{N}_Y(h)} (t - y_i)^+\right] \leq 0 \quad \forall t \in [0, \bar{z}_h], \forall h \in \mathcal{H}. \tag{D2}$$

We can state the following Lemma which corresponds to Lemma 1.

Lemma 3 The following condition is equivalent to Condition (D2).

$$\begin{aligned} &\sum_{i \in \widetilde{N}_X(h)} (t - \min\{x_i, \bar{z}_h\})^+ - \sum_{i \in \widetilde{N}_Y(h)} (t - \min\{x_i, \bar{z}_h\})^+ - \left(\widetilde{N}_X(h) - \widetilde{N}_Y(h)\right)(t - \bar{z}_h)^+ \\ &\qquad \leq 0, \ \ \forall t \in [0, \infty), \forall h \in \mathcal{H}. \end{aligned} \tag{D3}$$

Proof Noting that $t - \min\{x, \bar{z}_h\} = (x - \bar{z}_h)^+ + t - x$, we obtain $(t - \min\{x, \bar{z}_h\})^+ = (t - x)^+$ $\forall t \leq \bar{z}_h$. Thus, (D2) and (D3) are equivalent for $\forall t \in [0, \bar{z}_h]$. For the case of $t > \bar{z}_h$, (D3) can be represented as follows:

$$\sum_{i\in\widetilde{N}_X(h)}(t-\min\{x_i,\bar{z}_h\})^+ - \widetilde{N}_X(h)(t-\bar{z}_h)^+ = \sum_{i\in\widetilde{N}_X(h)}(\bar{z}_h - x_i)^+ \ \ \forall t \in (\bar{z}_h,\infty),$$

which implies (D2). ■

When we focus on the poverty measurement, it is sufficient to consider the income below the corresponding poverty line. For $\tilde{\mathbf{x}}_{\uparrow}^{h}$, let $\tilde{\mathbf{x}}_{\uparrow}^{h*}$ be a $\widetilde{N}_X(h)$-dimensional vector whose entries consist of the censored income by the poverty line: $x_{(i)}^{h*} \coloneqq \min\{x_{(i)},\bar{z}_h\}$ for $i\in\widetilde{\mathcal{N}}_X(h)$: $\tilde{\mathbf{x}}_{\uparrow}^{h*} \equiv (x_{(1)}^{h*},\dots,x_{(\widetilde{N}_X(h))}^{h*})$. In a similar way, we define $\tilde{\mathbf{y}}_{\uparrow}^{h*} \equiv (\tilde{y}_{(i)}^{h*},\dots,\tilde{y}_{(\widetilde{N}_Y(h))}^{h*})$.

To compare the distributions $\tilde{\mathbf{x}}^{*h}$ and $\tilde{\mathbf{y}}^{*h}$, we define a modified sub-vector,

$$\check{\mathbf{x}}^{h} := \begin{cases}[\tilde{\mathbf{x}}_{\uparrow}^{h*},\bar{\mathbf{z}}^{h}] & if\ \widetilde{N}_Y(h) > \widetilde{N}_X(h),\\ \tilde{\mathbf{x}}_{\uparrow}^{h*} & if\ \widetilde{N}_Y(h) \le \widetilde{N}_X(h),\end{cases} \tag{6a}$$

where $\bar{\mathbf{z}}^{h} \equiv (\bar{z}_h,\dots,\bar{z}_h)$ is an $|\widetilde{N}_X(h)-\widetilde{N}_Y(h)|$-dimensional vector in which all of the entries are equal to the poverty line for the h-th group. From the definition, $\check{x}_j^{h} = x_{(j)}^{h*}$ for $j\in\widetilde{\mathcal{N}}_X(h)$ and $\check{x}_j^{h} = \bar{z}_h$ for $j = \widetilde{N}_X(h)+1,\dots,\widetilde{N}_Y(h)$ if $\widetilde{N}_Y(h) > \widetilde{N}_X(h)$. Similarly, we define

$$\check{\mathbf{y}}^{h} = \begin{cases}[\tilde{\mathbf{y}}_{\uparrow}^{h*},\bar{\mathbf{z}}^{h}] & if\ \widetilde{N}_X(h) > \widetilde{N}_Y(h),\\ \tilde{\mathbf{y}}_{\uparrow}^{h*} & if\ \widetilde{N}_X(h) \le \widetilde{N}_Y(h).\end{cases} \tag{6b}$$

It should be noted that $\bar{z}_h$ is the maximum value of the elements of $\check{\mathbf{x}}^{h}$ and $\check{\mathbf{y}}^{h}$. Thus, we can apply Lemma 1 and obtain the similar results for the welfare dominance.

Lemma 4. Condition (D3) is equivalent to the following condition.

$$\sum_{j=1}^{r}\check{x}_j^{h} \ge \sum_{j=1}^{r}\check{y}_j^{h}, \qquad r = 1,\dots,\widehat{N}(h), \forall h \in \mathcal{H}. \tag{D4}$$

Proof. See Marshall et al. (2011, Proposition B.4, p.158).

Now, we can state the poverty dominance condition based on the GL curves comparison.

Proposition 2. The following two conditions are equivalent.

$$P_X \le P_Y \ \ \forall P \in \mathcal{P}_{CM}(\bar{\mathbf{z}}), \bar{\mathbf{z}} \in \mathcal{Z}, \tag{C}$$

$$GL(\check{\mathbf{x}}^h, p) \geq GL(\check{\mathbf{y}}^h, p) \;\; \forall p \in [0,1], \forall h \in \mathcal{H}. \tag{D5}$$

Proof. (D4)⇔(D5) is confirmed by a similar way in Proposition 1. From Lemmas 3 and 4, we know (C)⇔(D4).

The procedure for checking poverty dominance is straightforward. First, we draw the GL curve for each subgroup. Next, we modify the GL curves according to a poverty line that is set by researchers. Finally, by sequentially comparing the modified GL curves, we obtain insights into the poverty dominance.

Example 2 Suppose that we have the following distributions of income partitioned into three subgroups.

$$\mathbf{x}^1 = (2, 5, 14), \qquad \mathbf{x}^2 = (4, 5, 6, 10), \qquad \mathbf{x}^3 = (3, 7, 20),$$

$$\mathbf{y}^1 = (1, 3, 4, 10), \qquad \mathbf{y}^2 = (12, 15), \qquad \mathbf{y}^3 = (2, 5, 7, 8).$$

In this example, for any poverty line $\boldsymbol{z} \in \mathcal{Z}$, we can easily confirm that $GL(\check{\mathbf{x}}^h, p) \geq GL(\check{\mathbf{y}}^h, p)$, $\forall p \in [0,1]$ holds for $h \in \{1,3\}$. However, the dominance relation between $\check{\mathbf{x}}^2$ and $\check{\mathbf{y}}^2$ is less clear, since it depends on the poverty line. Figure 2 (*a*) shows the cumulative income for each distribution. If we consider a relative low poverty line such as $\bar{z}_2 = 6$, then condition (D5) holds, as shown in Figure 2 (*b*): this implies that $P_X - P_Y \geq 0$ for all $P \in \mathcal{P}_{CM}(\bar{\mathbf{z}})$ and the poverty lines $\bar{\mathbf{z}} = (\bar{z}_1, \bar{z}_2, \bar{z}_3)$, such as $\bar{z}_1 \geq \bar{z}_2 = 6 \geq \bar{z}_3$. In contrast, when we consider a relatively high poverty line such as $\bar{z}_2 = 12$, the two modified GL curves cross, as shown by the dotted lines in Figure 2 (*c*). This implies that condition (D5) does not satisfied, and thus there exists a poverty index $P \in \mathcal{P}_{CM}(\bar{\mathbf{z}})$ for $\bar{z}_1 \geq \bar{z}_2 = 12 \geq \bar{z}_3$ such that $P_X - P_Y < 0$.

[Figure 2 placed here]

5 Empirical Illustrations

In this section, we apply the procedure described in the previous sections to the income distributions of Italian households in 2006 and 2012. The data used for this illustration come from the Survey on Household Income and Wealth (SHIW) conducted by the Bank of Italy. We employed the data described in the Historical Database (version 8.0) of the SHIW. The sample size is 7768 in the 2006 survey and 8151 in the 2012 survey.[11]

Our analysis is based on the net disposable income (excluding income from financial assets), which is denoted as Y1 in the SHIW. The income unit used in the analysis is a household, that is, the cumulative number of income units is measured by household. In the dataset, there are a few households whose incomes are negative; we replaced these with zero income.

We partitioned the total sample into five subgroups, based on the number of members of each household. This may be too simple to capture the heterogeneity of households. However, we used this simple classification scheme since the equivalence scale recently used by OECD is also based on only household size.[12] The subgroups consist of one, two, three, four, and five or more persons. Based on the literature, we assumed that need increases with the number of household members: a household with five or more persons thus belongs to the first subgroup.

The descriptive statistics are summarized in Table 1. All of these statistics were calculated using the microdata adjusted by the sampling weight.[13] The consumer price index (CPI) that had a base year of 2005 was used for transforming the nominal income into the real income.[14] From Table 1, we can see that the household size decreased: the proportion of single households considerably increased while those of others decreased, with the exception of households of five or more persons. We can verify that the average income in 2012 was smaller than that in 2006. That is, the income distribution seen in 2012 was never dominated by that of 2006 in the sense of the extended SGL.

[Table 1 placed here]

Before presenting the investigation using the extended SGL criterion, let us consider the GL curve based on equivalent incomes. We will use the same equivalence scale as that used by the OECD: equivalent income is defined to be the income divided by the square root of the number of household members. Figure 1 shows the GL curves for 2006 and 2012 based on the equivalent income: note that the income distribution of 2006 dominates that of 2012 in the sense of the GL criterion. However, note that this result depends on the particular equivalence scale that is used.

---

[11] When we wish to compare distributions that have populations of different sizes, our procedure can be easily extended by constructing replicated distributions that have populations of same size. In this situation, the social evaluation function must have replication invariance, as is usually assumed for GL comparisons with populations of different size.

[12] For example, see OECD (2012).

[13] We used the PESOFIT in the SHIW as the sampling weight.

[14] The CPIs were 102.22 for 2006 and 117.53 for 2012.

[Figure 3 placed here]

We turn now to the welfare dominance in line with the extended SGL criterion. First, we must determine the income levels that result in the same utility for the different subgroups. Since our main goal is to illustrate the extended SGL criterion, we simply set as $\bar{v} = 788500$ euro; this is approximately the maximum level of income in the data for 2006 and 2012.

The GL curves for various subgroups are presented in Figure 4. From this figure, we can confirm that the extended SGL condition is violated: two distributions are not rankable. In particular, except for the neediest group, shown in Fig. 4 (a), and distribution of all data, shown in Fig. 4 (e), the GL curves cross. Thus, if we partition the households as those with five or more persons and all other households, the distribution for 2012 is dominated by that of 2006 according to the extended SGL criterion. However, when we allow for any differences between the various types of households, the dominance relations are not clear.

[Figure 4 placed here]

Next, we analyze the poverty gap. We set the poverty line for single households at 11250 euro, based on the 2005 consumer price: $\bar{z}_5 = 11250$. This poverty line is approximately equal to the average per capita income in 2006 and 2012. Based on this poverty line for single-person households, we set $\bar{z}_h = 11250 \times (6 - h)$ for $h \in \{1,2,3,4\}$.

Figure 5 shows the modified GL curves based on $\check{\mathbf{x}}^h$ and $\check{\mathbf{y}}^h$ for $h \in \{1,2,3,4\}$. In the analysis of welfare dominance, we already saw that the distribution of all of the data for 2012 is dominated by that of 2006 in the sense of GL. This implies that $GL(\check{\mathbf{x}}^5, p) \geq GL(\check{\mathbf{y}}^5, p)$, $\forall p \in [0,1]$ holds for any poverty line $\bar{z}_5$. Therefore, we present the GL curves for $j = 1, \dots, 4$. We can see that every modified GL curve for 2012 is dominated by that of 2006. That is, the distribution of 2012 is poverty dominated by that of 2006.

[Figure 5 placed here]

5 Discussion

In this paper, we propose a simple condition for the extended SGL that was proposed by Jenkins and Lambert (1993) and Chambaz and Maurin (1998). The procedure presented here can be easily implemented by comparison of the GL curves. This result is not surprising. It is well-known that the second-order stochastic dominance condition and its corresponding inverse condition are equivalent. In addition, the second order inverse function coincides with the GL function. Since the extended SGL condition is characterized by the second-order stochastic dominance condition, the condition is also represented by using the GL curve. This fact suggests that the dominance conditions other than the extended SGL can be represented by using the GL curves as long as the condition is based on the second-order stochastic dominance.

Moyes (2012) pointed out that the result of the comparisons may be changed by the choice of the upper bound of the income applied to the distribution functions. This difficulty is caused by assumption U2. Indeed, Moyes (2012) considered the following condition UM instead of U2.

UM $$U(y,1) \leq U(y,2) \leq, \dots, \leq U(y,H), \forall y \leq \bar{v}.$$

As mentioned in Section 2, conditions U1 and U2 imply condition UM. According to Moyes (2012), we introduce the following class of utility function:

$$\mathcal{U}_M = \{U \in \mathcal{U}: \text{U1, UM and U3 are satisfied}\}$$

Moyes (2012) has established that $W_X \geq W_Y \ \forall U \in \mathcal{U}_M$ is equivalent to the following condition BM (Moyes, 2012. Proposition 3.2):

$$F_X(\bar{v},h) \leq F_Y(\bar{v},h) \quad \forall h \in \mathcal{H} \backslash \{H\}. \tag{BM1}$$

$$\int_0^t [F_X(\xi,h) - F_Y(\xi,h)]d\xi \leq 0 \quad \forall t \leq \bar{v}, \forall h \in \mathcal{H}. \tag{BM2}$$

While condition (BM2) is the same as condition (B'), condition (BM1) requires that $\tilde{N}_X(h) \leq \tilde{N}_Y(h) \ \forall h \in \mathcal{H} \backslash \{H\}$. Fleurbaey et al. (2003) argued that Jenkins' and Lambet's criterion coincides with Moyes' one when $\bar{v}$ approaches to infinity[15]. It is possible to compare the distributions based on the GL curves presented in section 2. In Figure 1, we can test the Moye's criterion by comparing two solid lines. In this case, Moyes' criterion does not hold because $F_X(\bar{v},2) > F_Y(\bar{v},2)$.

We also argued that the poverty gap dominance can be tested by using the modified GL curves. Chambaz and Maurin (1998) showed that $P_X \leq P_Y \ \forall P \in \mathcal{P}_{CM}(\bar{\mathbf{z}}), \bar{\mathbf{z}} \in \mathcal{Z}$ is equivalent to $P_X \leq P_Y \ \forall P \in \mathcal{P}_{CM}(\bar{\mathbf{z}}')$, $\forall \bar{\mathbf{z}}' \leq \bar{\mathbf{z}}$, $\bar{\mathbf{z}}' \in \mathcal{Z}$. In the GL curve comparison presented in here, since $f_h(x) = z_h' - (z_h' - x)^+ = \min\{z_h', x\}$ is increasing concave in x, $\sum_{j=1}^r \check{x}_j^h \geq \sum_{j=1}^r \check{y}_j^h$ for $r = 1, \dots, \widehat{N}(h)$, $\forall h \in \mathcal{H}$ implies $\sum_{j=1}^r f(\check{x}_j^h) \geq \sum_{j=1}^r f(\check{y}_j^h)$ for $r = 1, \dots, \widehat{N}(h)$, $\forall h \in \mathcal{H}$ (See Marshall et al. 2011. Theorem A.2., p.167). Noting that $\bar{\mathbf{z}}' \leq \bar{\mathbf{z}}$ and $\bar{\mathbf{z}}' \in \mathcal{Z}$, we can confirm

---

[15] See Remark 5.1 in Fleurbaey et al.(2003).

that (D4) holds for the distribution generated by $\bar{\mathbf{z}}'$. Thus, (D5) for $\bar{\mathbf{z}} \in \mathcal{Z}$ is equivalent to $L(\check{\mathbf{x}}'^h, p) \geq GL(\check{\mathbf{y}}'^h, p)$ $\forall \bar{\mathbf{z}}' \leq \bar{\mathbf{z}}$, $\bar{\mathbf{z}}' \in \mathcal{Z}$.

As stressed in Foster and Shorrocks (1988), welfare and poverty dominance criteria share the common feature of stochastic dominance. This linkage is also observed in the present analysis: the extended SGL condition for both welfare and poverty dominances can be tested by the GL curves. Furthermore, it should be noted that if the poverty index $P(y, z)$ belongs to $\mathcal{P}_{\mathrm{CM}}(\mathbf{z}^*)$ and $\mathbf{z}^* = (\bar{v}, \dots, \bar{v})$, then $-P(y, z) \in \mathcal{U}_{JL}$. This feature suggests that the maximum conceivable income imposed in the welfare dominance condition can be flexibly set to the different sub-groups. Although such generalization may not be applicable as long as the difference in needs originates from the number of household members, it is expected to extend the applicability of the SGL criterion by incorporating different demographic condition to each sub-group.

References

Atkinson, A.B., “Measuring Poverty and Differences in Family Composition,” *Economica*, 59, 1-16, 1992.

Atkinson, A.B., and F. Bourguignon, "Income Distribution and Differences in Needs," in G.R. Feiwel, ed., *Arrow and the Foundations of the Theory of Economic Policy*, 350-370, Macmillan, London, 1987.

Bank of Italy, *Survey on Household Income and Wealth*, (http://www.bancaditalia.it.) Accessed 14 May 2014.

Chambaz, C., and E. Maurin, "Atkinson and Bourguignon's Dominance Criteria: Extended and Applied to the Measurement of Poverty in France," *Review of Income and Wealth*, 44, 497-513, 1998.

Ebert, U., “Sequential Generalized Lorenz Dominance and Transfer Principle,” *Bulletin of Economic Research*, 52, 113-122, 2000.

Fishburn, P.C. and R.G. Vickson, “Theoretical foundations of stochastic dominance,” in: G.A. Whitmore, M.C. Findlay (Eds.), *Stochastic Dominance: An Approach to Decision Making under Risk*, Heath, Lexington, MA, 1978, pp. 37–113.

Fleurbaey, M., C. Hagneré, and A. Trannoy, “Welfare comparisons with bounded equivalence scales,” *Journal of Economic Theory*, 110, 309–336, 2003.

Foster, J.E. and A. Shorrocks, “Poverty orderings and welfare dominance,” *Social Choice Welfare*, 5, 179–198, 1988.

Gastwirth, J.L. “A general definition of the Lorenz curve,” *Econometrica*, 39, 1037–1039, 1971.

Jenkins, S., and P. Lambert, "Ranking income distributions when needs differ," *Review of Income and Wealth*, 39, 337-356, 1993.

Lambert, P.J., *The Distribution and Redistribution of Income: A Mathematical Analysis*, 3rd ed., Manchester University Press, Manchester, 2001.

Lambert, P.J., and X. Ramos, “Welfare Comparisons: Sequential Procedures for Heterogeneous Populations,” *Economica*, 69, 549-562, 2002.

López-Laborda, J., and J. Onrubia, "Personal Income Tax Decentralization, Inequality, and Social Welfare," *Public Finance Review*, 33, 213-235, 2005.

Marshall, A.W., I. Olkin, and B.C. Arnold, *Inequalities: Theory of Majorization and its Applications*, 2nd. ed., Springer, New York, 2011.

Moyes, P., “Comparisons of heterogeneous distributions and dominance criteria,” *Journal of Economic Theory*, 147, 1351–1383, 2012.

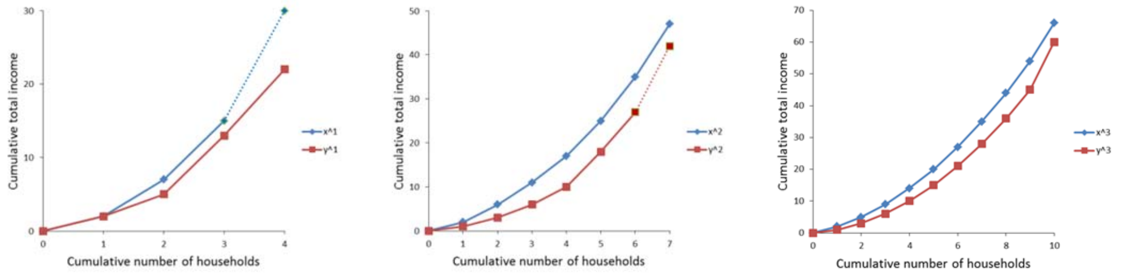
Muliere, P. and M. Scarsini, “A note on stochastic dominance and inequality measures,” *Journal of Economic Theory*, 49, 314-323, 1989.

OECD, *Quality Review of the OECD Database on Household Incomes and Poverty and the OECD Earnings Database Part I*, (http://www.oecd.org/els/soc/income-distribution-database.htm), 2012.

Ok, E.A., and P.J. Lambert, “On Evaluating Social Welfare by Sequential Generalized Lorenz Dominance,” *Economics Letters*, 63, 45-53, 1999.

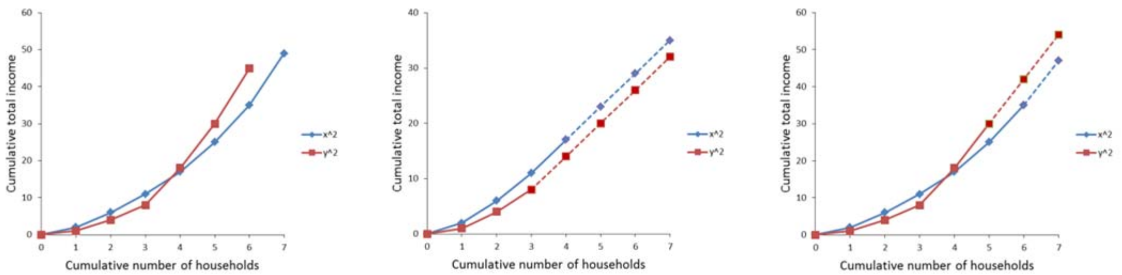
Ooghe, E., “Sequential Dominance and Weighted Utilitarianism,” *Economics Letters*, 94, 208-212, 2007.

Shorrocks, A. F., "Ranking Income Distributions," *Economica*, 50, 1-17, 1983.



(*a*) (*b*) (*c*)

Figure 1 Comparisons of GL curves for welfare dominance



$\bar{z}^2 = 6$ $\bar{z}^2 = 12$

(*a*) (*b*) (*c*)

Figure 2 Comparisons of GL curves for poverty dominance

Table 1 Distribution of Income in the SHIW

Number of household members	Share of subgroups		Mean (Euro)		Max. (Euro)		Min. (Euro)		C.V.	
	2006	2012	2006	2012	2006	2012	2006	2012	2006	2012
1	24.91	28.29	19474	15736	788495	185655	0	0	1.650	1.174
2	28.42	27.93	29583	26215	462111	231651	-8511	-851	1.194	1.203
3	21.52	19.38	35873	32499	385942	273122	0	0	1.156	1.218
4	18.48	17.66	39449	31714	637840	189526	0	0	1.345	1.173
5 or more	6.67	6.73	41156	31362	524359	185110	704	0	1.387	1.219
Total	100.00	100.00	31014	25786	788495	273122	-8511	-851	1.337	1.239

Source: Author's calculations from the SHIW historical data (version 8).
C.V. is the coefficient of variation. All income variables are deflated by the CPI (2005=100).

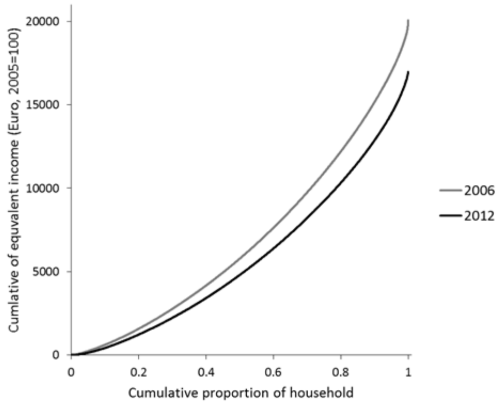
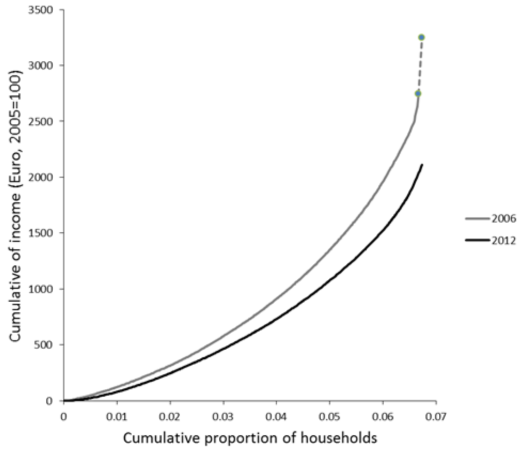
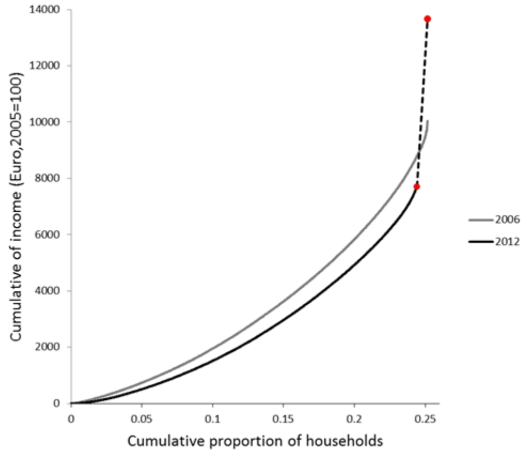


Figure 3 Comparison of GL curves based on equivalent income

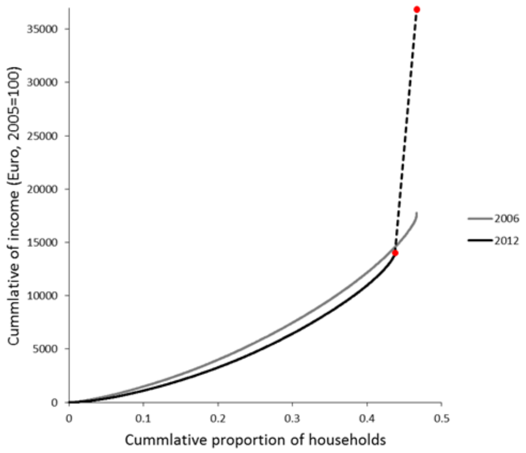
Source: Author’s calculation based on the SHIW historical data.



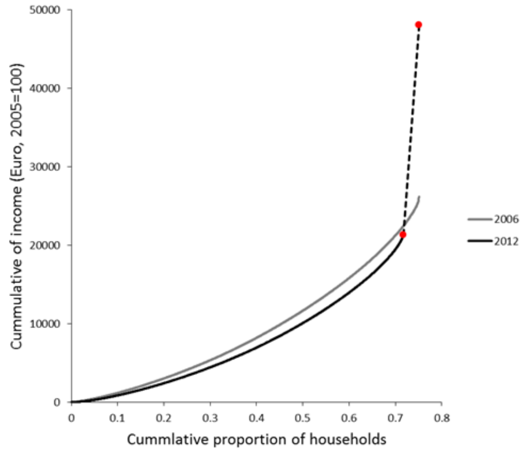
(*a*) Five or more household members



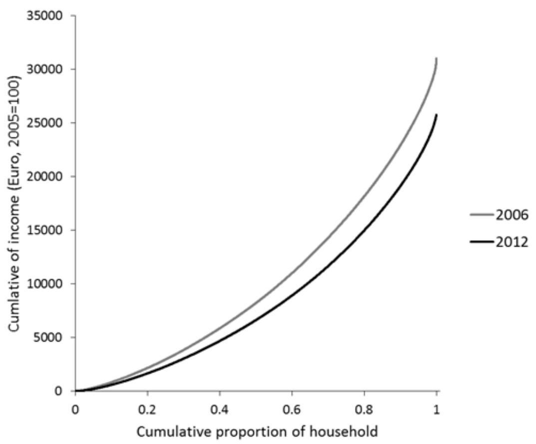
(*b*) Four or more household members



(*c*) Three or more household members



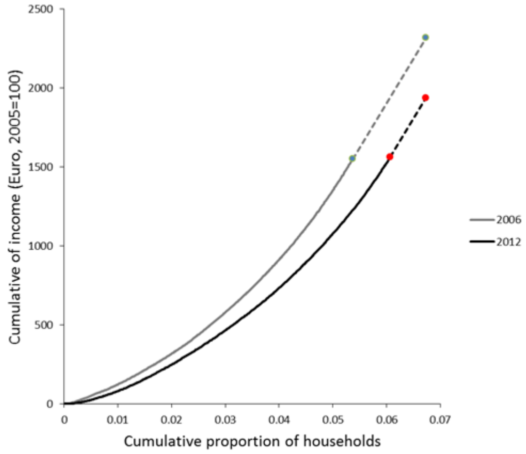
(*d*) Two or more household members



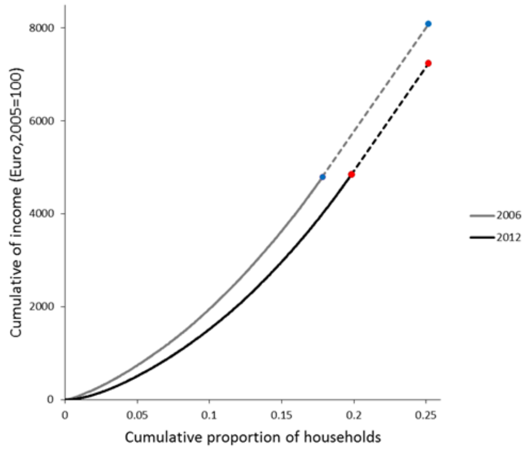
(*e*) All households

Figure 4 Welfare comparisons using the modified GL curves

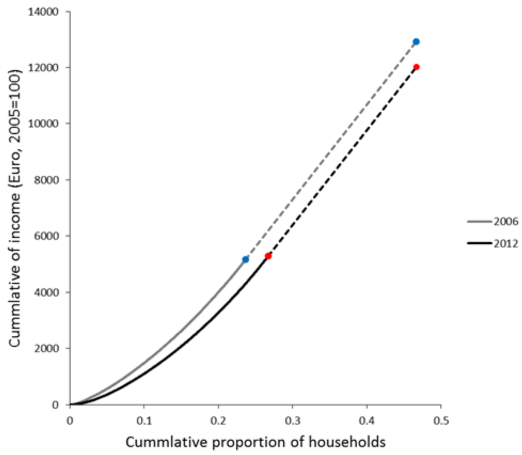
Source: Author's calculation based on the SHIW historical data.



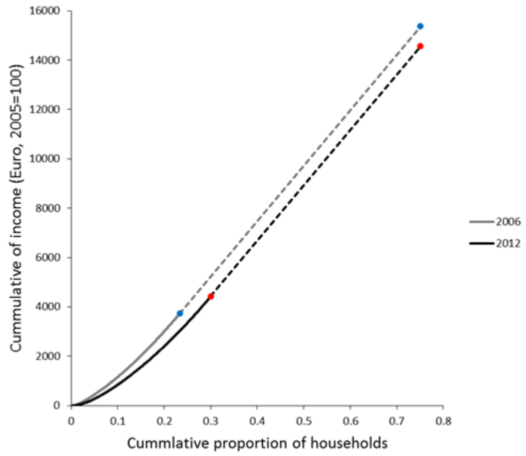
(*a*) Five or more household members



(*b*) Four or more household members



(*c*) Three or more household members



(*d*) Two or more household members

Figure 5 Poverty dominance evaluated by using the modified GL curves

Source: Author's calculation based on the SHIW historical data.