VYSOKÁ ŠKOLA BÁŇSKÁ
TECHNICKÁ UNIVERZITA OSTRAVA
EKONOMICKÁ FAKULTA

Central European Review of Economic Issues
EKONOMICKÁ REVUE

Parametric Families for the Lorenz Curve: An Analysis of Income Distribution in European Countries

Tommaso LANDO[*a], Michaela STANÍČKOVÁ[b], Jiří FRANEK[c]

[a]*Department of Management, Economics and Quantitative Methods, University of Bergamo (Italy) and Department of Finance, Faculty of Economics, VŠB – Technical University of Ostrava, Sokolská třída 33, 702 00 Ostrava, Czech Republic.*
[b]*Department of European Integration, Faculty of Economics, VŠB – Technical University of Ostrava, Sokolská třída 33, 702 00 Ostrava, Czech Republic.*
[c]*Department of Business Administration, Faculty of Economics, VŠB – Technical University of Ostrava, Sokolská třída 33, 702 00 Ostrava, Czech Republic.*

Abstract

The European Union Survey on Income and Living Conditions (EU-SILC) is the main source of information about living standards and poverty in the EU member states. We compare different parametric models for the Lorenz curve (LC) with an empirical analysis of the income distributions of 26 European countries in the year 2017. The objective of our empirical study is to verify whether simple mono-parametric models for the LCs can represent similarities or differences between European income distributions in sufficient detail, or whether an alternative, more sophisticated multi-parametric model should be used instead. In particular, we consider the power LC, the Pareto LC, the Lamè LC, a generalised bi-parametric version of the Lamè LC, a bi-parametric mixture of power LCs and the recently introduced arctan family of LCs. Whilst the first three families are ordered, in that different parametric values correspond to a situation of Lorenz ordering, the latter three may also identify the ambiguous situation of intersecting LCs. Therefore, besides focusing on the goodness-of-fit of the models considered and their mathematical simplicity, we evaluate the effectiveness of multi-parametric models in identifying the non-dominated cases.

Keywords

European Union, income inequality, Lorenz ordering, stochastic dominance.

JEL Classification: C44, E24, O52

---

[*] tommaso.lando@unibg.it
The research was supported through the *Czech Science Foundation (GACR) under project 17-23411Y.*

ISSN 1212-3951 (Print), 1805-9481 (Online)
ER-CEREI, Volume 21: 51–60 (2018).
doi: 10.7327/cerei.2018.06.03

Parametric Families for the Lorenz Curve: An Analysis of Income Distribution in European Countries

Tommaso LANDO, Michaela STANÍČKOVÁ, Jiří FRANEK

1. Introduction

The reduction of disparities in the European Union (EU) has become an important goal since its latest enlargements in 2004, 2007 and most recently in 2013. The accession of new members, each with less than half the GDP per capita of original EU member states (before 2004), has brought new challenges for a reinforced cohesion policy. Such a policy takes as its inspiration Article 130a of the Treaty on European Union and intends *to reduce disparities between the levels of development of the various countries and especially regions, and the backwardness of the least favoured countries, resp. regions including rural areas* (European Commission, 1996).

1.1 Economic background

At present, more than of two thirds of the European Structural and Investment Funds budget is allocated to countries with regions in which GDP per capita lags behind the EU average. Disparities have been extensively investigated in the economic literature by considering GDP per capita as a measure of disparity. The effectiveness of interventionist regional policies has often been evaluated in terms of convergence or divergence of per capita income across countries and specific regions, eventually giving each region a weight proportional to its population size (see, among others, Barro and Sala-i-Martin, 1991, 1995; Quah, 1996; Le Gallo, 2004; Pittau and Zelli, 2006; Sala-I-Martin, 2006; Allmendinger and Driesch, 2014). Findings are contingent upon the time span examined, the number of countries and hence number of regions, the level of disaggregation and the statistical method used. There is widespread agreement that income disparities across European regions belonging to the EU15[1] have narrowed over time, but the reduction of income disparities across regions cannot be equated with a reduction in disparities within regions. That is, a region with high GDP per capita may have substantial pockets of poverty, and a region with low GDP per capita may have some areas of prosperity. The directives of the European Commission implicitly assume not only that the funding received by a region will be converted to greater prosperity on average, but also reduce existing disparities in the region (De Rynck and McAleavey, 2001). Resources awarded to a region whose average income level is low may simply result in additional well-paid jobs for the narrow upper-middle class and ultimately lead to greater inequality.

Inequality and growth are interlinked, but it is difficult to establish the direction of causality (Aghion and Howitt, 1988). Studies of the effect of growth on inequality traditionally refer to the hypothesis of Kuznets (1955). This states that economic inequality increases over time while an area is developing, until it reaches a certain level of per capita GDP. After that, inequality begins to decrease. At the same time, the level of inequality may affect, positively or negatively, economic growth via distinct channels: accumulation of physical and human capital, redistributive public policies and political and social uncertainty (Weil, 2005), as also highlighted by Longford et al. (2010).

The regional (that is, sub-national) dimension enriches the debate on growth and income inequality. Monitoring income inequality as well as other indicators related to personal income distribution within European countries is reliant on comparable and internationally harmonised estimates for the EU member states (Staníčková, 2017). Comparable data on personal income distribution at the national level are difficult to obtain. Focusing on the EU, the annual European Statistics on Income and Living Conditions Survey (EU-SILC) is a principal source of data regarding the socioeconomic conditions of individuals and households in the EU countries and their regions. The Survey is an-

[1] The EU15 comprised the following 15 counties: Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, Netherlands, Portugal, Spain, Sweden, and United Kingdom.

nual and has a longitudinal and cross-sectional component. Therefore, EU-SILC is a reference source for comparative statistics on income, poverty and social exclusion. The survey design seeks to obtain representative samples at both EU and country levels, as well as for several subgroups, such as sex, household size, household type and socio-economic group (EUROSTAT, 2018).

1.2 Representation of income: main issues and objectives of the study

The Lorenz curve (LC) is a basic tool for the representation of income inequality (Lorenz, 1905). By definition, the LC evaluated in $p \in [0,1]$ gives the proportion of the total wealth corresponding to the $p100\%$ poorer part of the population, where perfect equality is represented by a LC with equation $L(p) = p$. Consequently, if the LC of a distribution is uniformly higher than the LC of another, we may argue that the first distribution exhibits a lesser degree of inequality compared to the other. Such a criterion is generally referred to as the Lorenz order, or Lorenz dominance (LD). However, we may note that the LD is not a complete order, that is, it is possible to find pairs of intersecting LCs, where neither of the two dominates the other. In particular, many empirical studies have revealed that such ambiguous cases are extremely frequent in practice, that is, many pairs of distributions cannot be ranked based on the LD. In such cases, we can compare intersecting distributions by relying on weaker orders of inequality, such as the LD of the second degree, as studied by Aaberge (2009).

In this paper, we are primarily concerned with the issue of finding a parametric functional form to represent the LC, and to investigate the effectiveness of a model in identifying dominated and non-dominated pairs of LCs. In this regard, many different functional forms for the LC have been proposed in the literature. Some belong to mono-parametric families, which may be denoted as ordered families of LCs. The main property of such families is that the LD is fully characterised by the value of a unique parameter. This may be considered an advantage in terms of simplicity of interpretation, although multi-parametric families generally yield better performances in terms of goodness-of-fit. However, by construction, ordered families do not permit LCs to cross, and this lack of flexibility may be inappropriate in a number of cases.

In section 2 we study some different families of parametric LCs, namely the power Lorenz curve (PLC); the Pareto LC (PARLC); the Lamè class of LCs (LLC) (Sarabia et al., 2017), which actually consists of two slightly different formulas; a generalised version of the Lamè curve (GLLC) studied by Sarabia et al. (1999); the mixed power LC (MPLC), that is, a bi-parametric mixture of PLCs; and the arctan family of LCs, recently introduced by Gòmez-Dèniz (2016). Whilst the PLC, the PARLC and the LLC are mono-parametric ordered families, the GLLC, MPLC and arctan families depend on two parameters, where different parametric combinations may yield intersecting LCs. Therefore, the main objectives of this paper are:

- To study the suitability of these different families in modelling income distribution in terms of goodness-of-fit. Special attention is given to the advantages/disadvantages of mono-parametric models with respect to multi-parametric models;
- To study the usefulness of multi-parametric families in describing non-dominated situations (i.e., cases of intersecting LCs). In particular, a multi-parametric model should be able to identify most of the crossing pairs of LCs.

For this purpose, in section 3 we perform an empirical analysis of the LCs of 26 European countries in 2017. The data have been downloaded from Eurostat's database (EUROSTAT, 2018). Section 4 discusses the results and draws conclusions.

2. Theory and methods

We first introduce some basic notation and definitions. We recall that a preorder is a binary relation $\leq$ over a set S that is reflexive and transitive. In particular, observe that a preorder $\leq$ does not generally satisfy the antisymmetric property (that is, $a \leq b$ and $b \leq a$ does not necessarily imply $a = b$) and it is generally not complete (that is, each pair a, b in S is not necessarily related by $\leq$).

Let F be a non-negative distribution with positive and finite expectation μ_F. The (generalised) inverse or quantile function of F is given by

$$F^{-1}(p) = \inf\{z: F(z) \geq p\} p \in (0,1). \quad (1)$$

The Lorenz curve is an increasing and convex function $L_F: [0,1] \to [0,1]$ defined as follows (Gastwirth, 1971):

$$L_F(p) = \frac{1}{\mu_F}\int_0^p F^{-1}(t)dt, p \in (0,1). \quad (2)$$

We recall that the Gini index is given by twice the area between the Lorenz curve and the 45° line:

$$\Gamma(F) = 1 - 2\int_0^1 L_F(t)dt. \quad (3)$$

In fact, for a given percentage p, $L_F(p)$ represents the percentage of *total* possessed by the low $100p\%$ part of the distribution. It is well-known that the higher of two non-intersecting Lorenz curves can be obtained from the lower by a sequence of income transfers from *richer* to *poorer* individuals. This criterion has been called the *Pigou-Dalton condition*. For this reason, in an economic framework the higher of two non-intersecting LCs should be preferred, in that it shows less inequality compared with the lower. This idea defines the LD, that is, a preorder ($\leq_L$) in the space of non-negative distributions (with finite mean).

Definition 1: We write $F \leq_L G$ if and only if $L_F(p) \geq L_G(p), \forall\, p \in (0,1)$.

However, the LD is not a complete order, that is, it is possible to find pairs of LCs that are not ranked by the LD. Indeed, LCs may cross once or several times. Ranking such ambiguous situations represents a fundamental issue in the literature, in that is has been empirically shown that intersecting LCs constitute an extremely frequent situation in practice. To address this issue and obtain unambiguous rankings, some weaker criteria have been introduced in the literature. Muliere and Scarsini (1989) and Aaberge (2009) suggest cumulating LCs from the left or right, that is, attaching more weighting to low or top incomes. Lando and Bertoli-Barsotti (2016) suggest cumulating from both tails of the distribution towards the *centre*, that is, attaching more weighting to low and top incomes simultaneously.

In order to understand the relations between the models analysed in the sequel, we provide the definition of a *dual* LC. For a given LC L_F, we can compute the complementary LC

$$\begin{aligned}\bar{L}_F(p) &= \frac{1}{\mu_F}\int_0^p F^{-1}(1-t)dt \\ &= 1 - L_F(1-p), p\ (0,1)\end{aligned} \qquad (4)$$

that is, an increasing and concave curve $\bar{L}_F: [0,1] \to [0,1]$ which represents the proportion of total wealth corresponding to the top $p100\%$ richer part of the population. Whilst $\bar{L}_F$ cannot be considered a LC, its inverse function $\bar{L}_F^{-1}$ does. In particular,

$$\bar{L}_F^{-1}(p) = 1 - L_F^{-1}(1-p) \qquad (5)$$

represents the proportion of population that holds the top $100p\%$ part of the total wealth. We observe that $\bar{L}_F^{-1}$ is also a LC, in that it is non-decreasing, convex, differentiable almost everywhere and defined on the set $[0,1]$ $(\bar{L}_F^{-1}(0) = 0, \bar{L}_F^{-1}(1) = 1)$. From a geometrical point of view, the LC $\bar{L}_F^{-1}$ is symmetrical with respect to the LC L_F, where the axis of symmetry is the line $t = 1 - p$. We define $\bar{L}_F^{-1}$ as the dual version if the LC L_F.

Among the different functional forms that have been proposed to approximate the LCs of income distributions, we propose the following.

Definition 2: The power LC.

The PLC is defined by the following formula:

$$L_P(p,a) = p^a, p \in (0,1), a \geq 1. \qquad (6)$$

This basic model is clearly ordered with respect to the parameter a, as $L_P(p,a_1) \geq L_P(p,a_2)$ iff $a_1 \leq a_2$. For $a = 1$ we obtain the equality line.

Definition 3: The Pareto LC.

The PARLC is dual to the PLC and is defined by the following formula:

$$L_{PAR}(p,a) = 1 - L_P^{-1}(p,1/a) = 1 - (1 - p)^a, p \in (0,1), a \in (0,1]. \qquad (7)$$

Whilst the PLC is more suitable to represent inequality in the left tail, given its shape, the PARLC is more suitable to represent inequality in the right one.

This model is ordered with respect to the parameter a, as $L_{PAR}(p,a_1) \geq L_{PAR}(p,a_2)$ iff $a_1 \geq a_2$. For $a = 1$ we obtain the equality line.

Definition 4: The Lamè LC.

The LLC is closely related to the PARLC and is defined by the following two different formulas (LLC(1) and LLC(2), respectively):

$$L_{L1}(p,a) = L_{PAR}(p,a)^{1/a} = [1 - (1 - p)^a]^{1/a}, \quad p \in (0,1), a \in (0,1] \qquad (8)$$

$$L_{L2}(p,a) = 1 - (1 - p^a)^{1/a}, p \in (0,1), a \geq 1. \qquad (9)$$

In both cases, for $a = 1$ we obtain the equality line. The LLC has been introduced by Henle et al. (2008) and has more recently been studied by Sarabia et al. (2017). Both curves are ordered with respect to the parameter a, in particular:

$$L_{L1}(p,a_1) \leq L_{L1}(p,a_2) \text{ iff } a_1 \leq a_2,$$
$$L_{L2}(p,a_1) \leq L_{L2}(p,a_2) \text{ iff } a_1 \geq a_2.$$

Definition 5: The generalised Lamè LC.

The GLLC is defined by the following formula:

$$L_{GL}(p,a,b) = L_{PAR}(p,a)^b = [1 - (1 - p)^a]^b, \quad p \in (0,1), a \in (0,1], b \geq 1. \qquad (10)$$

For $a = b = 1$ we obtain the equality line. In contrast to L_P, L_{L1} and L_{L2}, this family is not ordered, in that different combinations of a and b may yield intersecting LCs.

In order to better-identify intersecting situations, we introduce another bi-parametric model.

Definition 6: The mixed power LC.

The MPLC is defined by the formula

$$L_M(p,a,b) = bp^a + (1 - b)(1 - (1 - p)^{1/a})\ b \in [0,1], a \geq 1. \qquad (11)$$

Then, L_M is simply a mixture of a PLC p^a and its dual version, i.e., the PARLC $1 - (1 - p)^{1/a}$. For $a = 1$ we obtain perfect equality, regardless of b. The mixing parameter b may determine a PLC ($b = 1$) or a PARLC ($b = 0$), and different combinations of bs may determine a large number of intersecting cases. We also note that the MPLC is closely related to the LLCs, as well as the PLC. Indeed, $L_{L2}(p,a) = 1 - (1 - p^a)^{1/a} = L_M(p^a,a,0)$. Therefore, we might also generalise the LLC with an alternative *mixed* approach by considering, for instance, $L_M(p^a, a, b)$. In so doing, for $b = 1$ we obtain the PLC, for $b = 0$ we obtain the LLC(2).

Definition 7: The arctan LC.

The arctan-LC, recently proposed by Gòmez-Dèniz (2016), is defined by the following formula

$$L_{ARC}(p, b) = 1 - \frac{\arctan(b(1-p))}{\arctan b}, b \in \mathbb{R} . \qquad (12)$$

Given that the composition of LCs yields a LC, for a given LC, say L, $L_{ARC}(L(p), b)$ is again a LC. In particular we consider $L_{ARC}(p^a, b)$, $L_{ARC}(L_{PAR}(p, a), b)$, $L_{ARC}(L_{L1}(p, a), b)$, namely the arctan-PLC, arctan-PARLC, arctan-LLC(1) and arctan-LLC(2).

In the next section, we compare the performance of these different parametric models in terms of goodness-of-fit, and we also analyse the effectiveness of the bi-parametric models in identifying non-dominated cases.

3. Empirical analysis

Many European countries have faced sluggish growth over the past decades and the trend has worsened in recent years. Contrary to economic booms, when most individuals are likely to see substantial increases in income, low growth tends to engender concerns about stagnating incomes, rising inequality and poverty. While inequality is often viewed from a national perspective, there are good reasons to analyse it for the EU as a whole.

Data have been retrieved online from Eurostat's website (the data have also been studied in Lando et al., 2017). In particular, Eurostat (EUROSTAT, 2018) provides *distributions of income by quantiles* with two options in terms of income and living conditions indicators, namely: i) *top cut-off point*, which represents the income of the individual at the right end of the given quantile and; ii) *share of national equalised income*, which is the share of the total income belonging to a given interval. Eurostat provides i) and ii) for the three quartiles, the four quintiles, the nine deciles and the first (and last) five percentiles. It should be stressed that some countries present negative incomes in the first 1–2 percentiles. Clearly, the presence of negative incomes contradicts the assumption of non-negativity of income distribution. Moreover, smaller percentile values are generally less reliable and accurate. Thus, we decided to consider the LCs, starting with $p = 0.03$. Indeed, by properly cumulating the shares of national equalised income we can obtain the values of the LC for

$$p = 0,0.03,0.04,0.05,0.1,0.2,0.25,0.3,0.4,0.5,$$
$$0.6,0.7,0.75,0.8,0.9,0.94,0.95,0.96,0.97,0.98,0.99,1,$$

that is, a LC with 20 nodes (excluding 0 and 1). We denote the set of 20 values above as S_{20}.

The first step of our analysis consists of analysing the empirical LCs computed from the observed data. For the year 2017, we obtained the LCs of 26 countries and compared each pair of LCs (i.e., 26×25/2 pairs) based on the LD relation. We find that the LD can rank only 44% of the pairs, while the remaining 56% of the pairs present intersecting LCs (i.e., with one or more crossings). Thus, LD seems not to be an effective criterion in comparing the LCs of European countries, because it cannot rank even half of the pairs. In particular, we find that 67% of crossing cases present one single crossing, 29% present two crossings and only 0.2% present more than two crossings.

As a second step, we then estimated the parametric models studied in section 2 by fitting them to the 26 observed LC discussed above. For the sake of simplicity, we used the least squares method, seeks to find the parametric values that minimise the square of the *distance* between the observed LC and the model considered. Such distance is actually a sum of quadratic differences, evaluated in all 20 nodes of the observed LCs. In particular, let L_{obs} be the empirical (observed) LC and let $L_{\boldsymbol{\alpha}}$ be the parametric LC, where $\boldsymbol{\alpha}$ is a parameter (possibly vectorial). We find the parametric value $\widehat{\boldsymbol{\alpha}}$ that minimises the distance.

$$SQE(\boldsymbol{\alpha}) = \sum_{p \in S_{20}} (L_{obs}(p) - L_{\boldsymbol{\alpha}}(p))^2. \qquad (13)$$

By dividing this distance by the number of nodes (20) and computing its square root, we obtain the *root mean squared error* (MSE), i.e. $RMSE(\widehat{\boldsymbol{\alpha}}) = \sqrt{SQE(\widehat{\boldsymbol{\alpha}})/20}$, which can be used as a measure of goodness-of-fit. We computed the RMSEs over the 26 LCs, for all the models considered, and obtained the following results. We also report the values of the Gini index, because the fit may be related to the shape of the LCs (logically, all the models would provide perfect fit to those LCs that correspond to the equality line, i.e., yield values of the Gini index equal to 0).

In Table 1, the results show that the LLC (in both versions) provide an excellent performance in approximating the LC, whilst the PLC and the PARLC are not well-fitting. The arctan family, in its different combinations, always provides a worse fit, besides having two parameters, compared to the mono-parametric LLCs. Therefore, we shall exclude the arctan family from the following discussion. The GLLC and the MPLC provide smaller RMSEs compared to the LLCs (obviously) but the improvement is questionable, especially if we are interested in the goodness-of-fit as well as the mathematical simplicity of the models considered. Indeed, the LLCs perform extremely well besides having just one parameter. Moreover, we compute the correlation between the RMSEs and the values of the Gini index. The results are: 0.57 for the PLC, 0.65 for the PARLC, 0.22 for the LLC(1), 0.25 for the LLC(2), 0.07 for the GLLC and 0.43 for the MPLC. Thus, the performance of the GLLC seems to be independent from the shape of the LC, unlike that of other models. In particular, the models related to the Lamè class seem to be less sensitive to the shape, whilst the power and Pareto-type models seem to be more suitable for LCs that show a more *even* distribution of income. The motivation can be found in the shape of the LCs considered. In particular, we found that income distribution in Europe is quite evenly distributed in the *centre*, whilst inequality is generally concentrated in the tails (i.e., low and high incomes); in Figure 1 displayed generally across the EU countries, and in Figures 2 and 3 for selected states. Therefore, the PLC may be inaccurate for representing the right tail, whilst the PARLC may be inaccurate for representing the left one. The goodness-of-fit analysis does not reveal whether the models considered are effective at identifying dominated or non-dominated pairs of LCs. As discussed in section 2, all of the ordered families cannot generate crossing LCs by construction. The next step of our study is focused simply on the GLLC and the MPLC. We compute the number of times when observed LCs cross as well as estimated LCs vs. the times when LCs do not cross but the estimated ones do, and so on. In particular, we find that the GLLC is able to identify the LD in 69% of cases whilst the PMLC in 74% (these percentages are computed over the number of cases when the LD holds). On the other hand, the percentage of cases when intersecting estimated LCs correspond to intersecting observed LCs (computed over the set of intersecting observed LCs) is 76% and 68% for the GLLC and the PMLC, respectively.

Table 1 RMSEs of the models considered and values of the Gini index

	PLC	*PARLC*	*LLC(1)*	*LLC(2)*	*GLLC*	*MPLC*	*arc-PLC*	*arc-PARLC*	*arc-LLC(1)*	*arc-LLC(2)*	*Gini*
BE	0.025	0.027	0.003	0.003	0.002	0.001	0.020	0.008	0.007	0.010	0.259
BG	0.055	0.030	0.012	0.010	0.006	0.005	0.047	0.013	0.016	0.022	0.399
CZ	0.032	0.018	0.006	0.006	0.002	0.002	0.025	0.006	0.011	0.015	0.245
DK	0.038	0.020	0.012	0.011	0.008	0.007	0.033	0.012	0.017	0.021	0.273
DE	0.033	0.026	0.004	0.003	0.003	0.002	0.027	0.009	0.009	0.013	0.289
EE	0.023	0.041	0.012	0.013	0.005	0.008	0.015	0.007	0.005	0.003	0.315
EL	0.034	0.033	0.004	0.003	0.004	0.003	0.028	0.011	0.008	0.013	0.330
ES	0.028	0.039	0.007	0.008	0.001	0.003	0.022	0.011	0.003	0.007	0.338
FR	0.040	0.021	0.010	0.009	0.004	0.003	0.033	0.009	0.014	0.019	0.290
HR	0.025	0.034	0.006	0.007	0.001	0.002	0.019	0.009	0.003	0.007	0.297
IT	0.031	0.034	0.004	0.004	0.003	0.004	0.026	0.012	0.007	0.011	0.325
CY	0.041	0.022	0.009	0.008	0.003	0.001	0.032	0.007	0.012	0.016	0.306
LV	0.030	0.039	0.006	0.007	0.002	0.005	0.023	0.008	0.002	0.006	0.344
LT	0.040	0.036	0.001	0.002	0.001	0.005	0.032	0.009	0.005	0.011	0.371
LU	0.032	0.031	0.001	0.002	0.001	0.003	0.025	0.008	0.006	0.010	0.309
HU	0.030	0.026	0.002	0.002	0.002	0.003	0.024	0.008	0.008	0.012	0.279
MT	0.032	0.026	0.002	0.002	0.001	0.003	0.024	0.006	0.007	0.011	0.283
NL	0.031	0.024	0.003	0.002	0.002	0.002	0.024	0.008	0.009	0.013	0.268
AT	0.030	0.026	0.003	0.002	0.003	0.003	0.025	0.009	0.009	0.013	0.277
PL	0.032	0.027	0.002	0.002	0.001	0.003	0.026	0.008	0.008	0.012	0.290
PT	0.039	0.030	0.003	0.003	0.001	0.004	0.031	0.008	0.008	0.014	0.333
RO	0.023	0.042	0.011	0.012	0.001	0.003	0.018	0.011	0.004	0.005	0.327
SI	0.022	0.025	0.002	0.002	0.000	0.002	0.017	0.006	0.005	0.008	0.235
SK	0.019	0.027	0.006	0.006	0.003	0.004	0.016	0.009	0.006	0.009	0.230
FI	0.032	0.019	0.006	0.006	0.002	0.001	0.025	0.006	0.011	0.014	0.253
SE	0.033	0.025	0.008	0.007	0.007	0.005	0.028	0.012	0.013	0.017	0.278
Mean	*0.032*	*0.029*	*0.006*	*0.005*	*0.003*	*0.003*	*0.025*	*0.009*	*0.008*	*0.012*	*0.298*

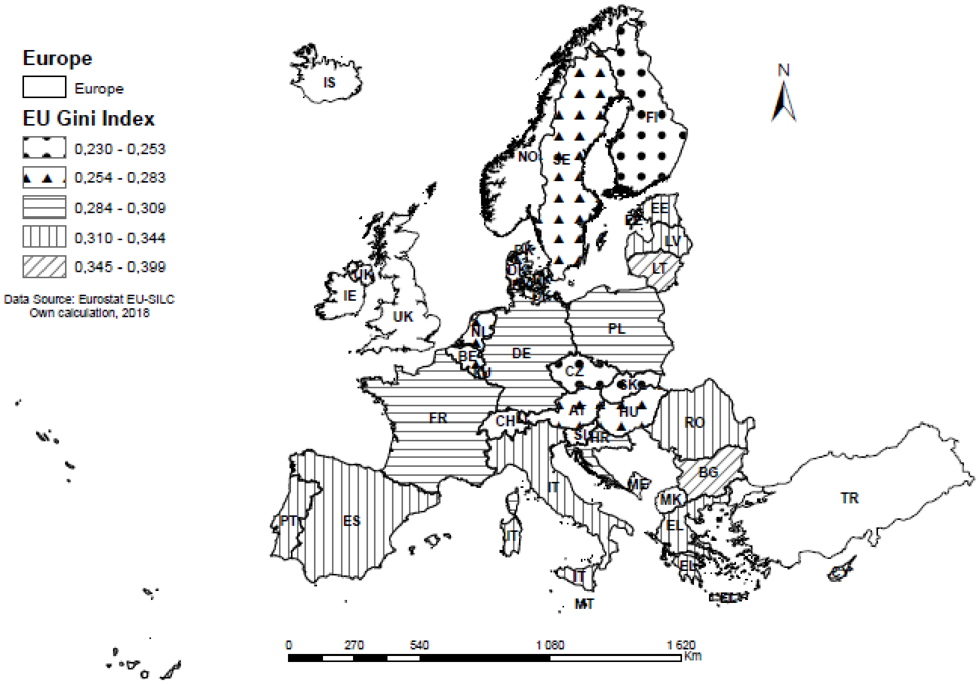

Figure 1 Geographical distribution of Gini index across the EU countries.

Thus, from these results, the PMLC seems to be more effective in identifying the dominated pairs whilst the GLLC seems to be more effective in identifying the non-dominated ones. We stress, however, that such percentages should be interpreted logically. Indeed, all of the ordered families of LCs would identify LD in 100% of the pairs in which LD is verified, just by construction. Therefore, given that the percentage of effectiveness in the non-dominated pairs is more reliable, we argue that the GLLC is preferable to the PMLC in terms of identification of the LD. The empirical results also reveal that inequality has been declining in most EU countries (Halásková et al., 2017).

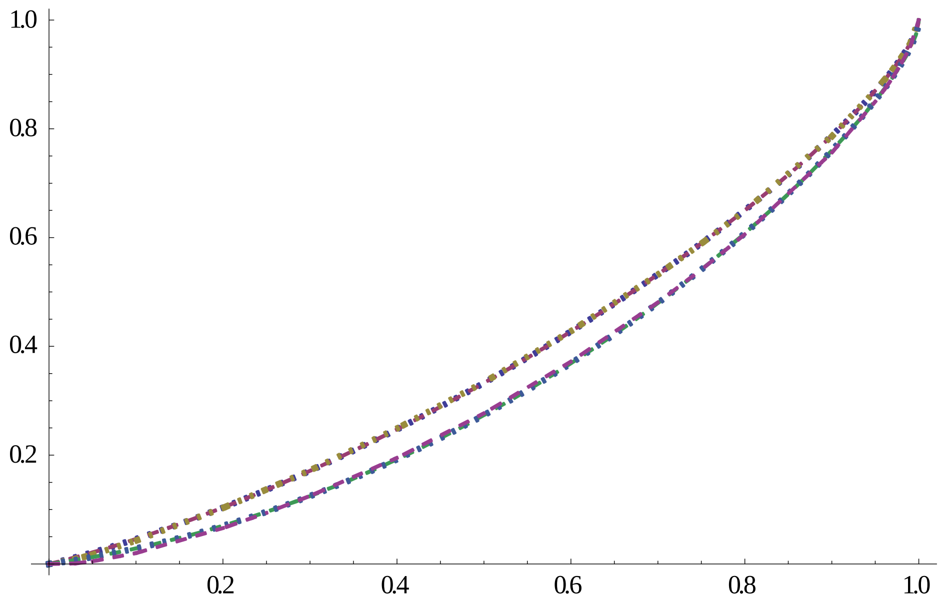

Figure 2 LC of the Czech Republic and Italy. Observed LCs (dashed line) vs. estimated LCs (MPLC, dotted line). The observed and estimated curves match almost perfectly, the LC of Czech Republic is never below that of Italy (LD is verified).

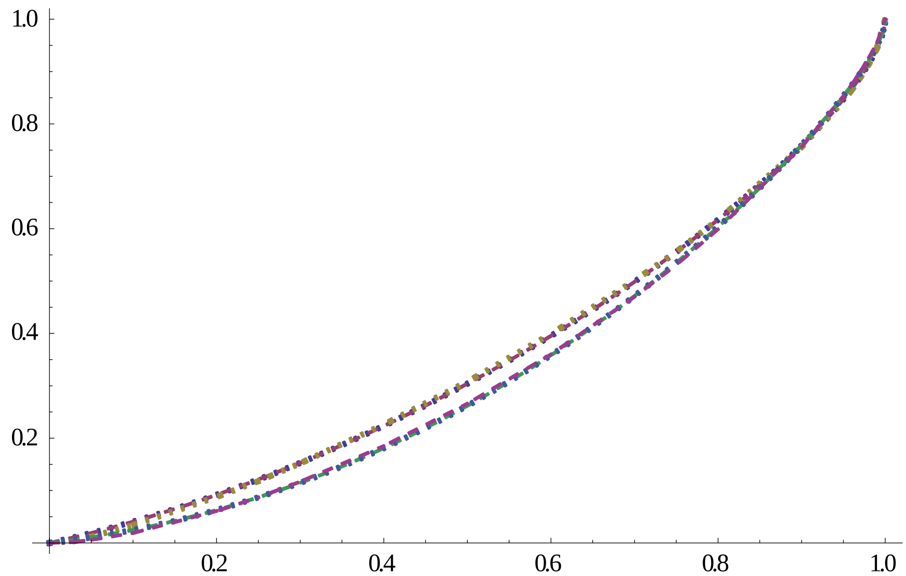

Figure 3 LC of Spain and France. Observed LCs (dashed line) vs. estimated LCs (MPLC, dotted line). The observed and estimated curves match almost perfectly; the LC of Spain starts below that of France and crosses in a neighbourhood of 0.9 (LD is not verified).

4. Conclusion

Our results highlight some of the advantages and disadvantages of ordered families of LCs compared to bi-parametric families. Ordered families are easy to interpret and may provide an extremely accurate approximation: this is the case, for instance, of the Lamè class, which seems to be able to capture the shape of an LC, at least with regard to the data set considered. Moreover, the unique parameter contains most of the information regarding the characteristics and concentration patterns of the population, and can therefore be used as an index of inequality, just like the Gini index. On the other hand, the two bi-parametric models examined generalise i) the Lamè class by introducing an additional power parameter; and ii) the power class, by introducing an additional mixing parameter. As an obvious consequence, the fit is enhanced, although this may be insufficient to justifying the use of a bi-parametric model. In this regard, the main advantage of a generalised model is the possibility of generating scenarios of crossing LCs, which happen to be quite frequent in practical situations. Indeed, our analysis shows that LCs of European countries cross more than half the time, hence it may be inappropriate to describe a partially ordered set of LCs with a totally ordered family. In contrast, the GLLC and the MPLC partially address this issue by supporting some of the cases when LCs do (or do not) cross. We conclude by noting that the LCs of the data set considered provide very similar shapes and concentration values, presenting most of their differences (of shape) in the tails.

Considering inequality in the European Union as a whole enables us to see what differences in income growth per decile has implied for inequality between individuals in the organisation, given that in parallel to changes in the distribution of income within countries, some economies have grown more rapidly than others. Second, we can no longer assume that all EU countries are comfortably on the downward part of the Kuznets curve, with inequality falling over time.

Poor growth performance in recent decades in Europe has crystallised concerns for rising income dispersion and social exclusion. European authorities have launched the Europe 2020 Strategy with the objective of reducing social inclusion in Europe on top of already existing European regional policies aimed at reducing regional disparities through stimulating growth in areas where incomes are relatively low. While it is most common to confine measures of inequality to national borders, the existence of such union-wide objectives and policies motivates measuring income dispersion among all Europeans in this paper. Towards the end of the 2000s, the income distribution in Europe was more unequal than in the average OECD country, albeit notably less so than in the United States. It is the within-country not between-country dimension that appears to be most important. Inequality in Europe has risen quite substantially since the mid-1980s. While the EU enlargement process has contributed to this, it is not the only explanation because inequality has also increased within a *core* of European countries. Large income gains among the 10% of top earners seem to constitute a primary driver behind this phenomenon.

References

AABERGE, R. (2009). Ranking intersecting Lorenz curves. *Social Choice and Welfare* 33: 235–259. https://doi.org/10.1007/s00355-008-0354-4

AGHION, P., HOWITT, P. (1998). *Endogenous growth theory*. Cambridge: MIT Press.

ALLMENDINGER, J., DRIESCH, E. (2014). Social inequalities in Europe: Facing the challenge. *Discussion Paper P* 2014–005. Berlin: Social Science Center.

BARRO, R., SALA-I-MARTIN, X. (1995*). Economic growth*. New York: McGraw Hill.

BARRO, R., SALA-I-MARTIN, X. (1991). Convergence across states and regions. *Brooking Papers on Economic Activity* 22, 107–182. https://doi.org/ 10.2307/2534639

DE RYNCK, S., MCALEAVEY P. (2001). The Cohesion deficit in Structural Fund Policy. *Journal of European Public Policy* 8: 541–557. https://doi.org/ 10.1080/13501760110064384

EUROPEAN COMMISSION. (1996). *First Report on Economic and Social Cohesion*. Luxembourg: Publication Office of the European Union.

GASTWIRTH, J.L. (1971). A general definition of the Lorenz curve. *Econometrica* 39: 1037–1039. https://doi.org/10.2307/1909675

GÓMEZ-DÉNIZ, E. (2016). A family of arctan Lorenz curves. *Empirical Economics* 51(3): 1215–1233. https://doi.org/10.1007/s00181-015-1031-y

HALÁSKOVÁ, R., MERIČKOVÁ MIKUŠOVÁ, B. (2017). Socio-economic development, income inequality and redistribution. *Ekonomická revue – Central European Review of Economic Issues* 20: 29–41. https://doi.org/10.7327/cerei.2017.03.03

HENLE, J.M., HORTON, N.J., JAKUS, S.J. (2008). Modelling inequality with a single parameter. In *Modeling income distributions and Lorenz curves*, 255–269. New York: Springer. https://doi.org/10.1007/978-0-387-72796-7_14

KUZNETS, S. (1955). Economic growth and income inequality. *American Economic Review* 45: 1–28. https://doi.org/10.2307/1811581

LANDO, T., BERTOLI-BARSOTTI, L. (2016). Weak orderings for intersecting Lorenz curves. *Metron* 74(2): 177–192. https://doi.org/10.1007/s40300-016-0087-6

LANDO, T., BERTOLI-BARSOTTI, L., STANÍČKOVÁ, M. (2017). Weak orders for Intersecting Lorenz Curves: Analysis of income distribution in European countries. In: *Proceedings of the 35th International Conference on Mathematical Methods in Economics*. University of Hradec Králové, 402–407.

LE GALLO, J. (2004). Space-time analysis of GDP disparities among European regions. *International Regional Science Review* 27: 138–163. https://doi.org/10.1177/0160017603262402

LONGFORD, N.T., PITTAU, M.G., ZELLI, R., MASSARI, R. (2010). Measures of poverty and inequality in the countries and regions of EU. *Working Paper Series – ECINEQ WP 2010 – 182*, 1–30. Rome: Society for the Study of Economic Inequality.

LORENZ, M.O. (1905). Methods of measuring the concentration of wealth. *Publications of the American Statistical Association*. 970: 209–219. https://doi.org/10.2307/2276207

MULIERE P., SCARSINI M. (1989). A note on stochastic dominance and inequality measures. *Journal of Economic Theory* 49(2): 314–323. https://doi.org/10.1016/0022-0531(89)90084-7

PITTAU, M.G., ZELLI, R. (2006). Empirical Evidence of Income Dynamic across EU Regions. *Journal of Applied Econometrics* 21: 605–628. https://doi.org/10.1002/jae.855

QUAH, D. (1996). Regional convergence clusters in Europe. *European Economic Review* 35: 951–958. https://doi.org/10.1023/A:1009781613339

SALA-I-MARTIN, X. (2006). The world distribution of income: Falling poverty and… convergence, period. *The Quarterly Journal of Economics* 121(2): 351–397. https://doi.org/10.1162/qjec.2006.121.2.351

SARABIA, J.M., CASTILLO, E., SLOTTJE, D.J. (1999). An ordered family of Lorenz curves. *Journal of Econometrics* 91(1): 43–60. https://doi.org/10.1016/ S0304-4076(98)00048-7

SARABIA, J.M., JORDÁ, V., TRUEBA, C. (2017). The Lamé class of Lorenz curves. *Communications in Statistics-Theory and Methods* 46(11): 5311–5326. https://doi.org/10.1080/03610926.2013.775306

STANÍČKOVÁ, M. (2017). Can the implementation of the Europe 2020 Strategy goals be efficient? The challenge for achieving social equality in the European Union. Equilibrium. *Quarterly Journal of Economics and Economic Policy* 12(3): 383–398. https://doi.org/10.24136/eq.v12i3.20

WEIL, D.N. (2005). *Economic Growth*. London: Pearson Education.

Additional resources

EUROSTAT (2018). *Mean and Median Income by Age and Sex - EU-SILC Survey*. [Online], accessed at 20. 10. 2018. Available from: <https://ec.europa.eu/eurostat/data/database>.