



Analysis of the inequality into distributions. An alternative approach to the Gini index applied to the spending environmental in EU

A. Seijas-Macias^{1,3}, A. Oliveira^{2,3} & T. A. Oliveira^{2,3}

¹*Department of Economy, Universidade da Coruña (Spain)*

²*Department of Science and Technology, Universidade Aberta (Portugal)*

³*CEAUL – Centro de Estatística e Aplicações da Universidade de Lisboa (Portugal)*

antonio.smacias@udc.gal, Amilcar.Oliveira@uab.pt, Teresa.Oliveira@uab.pt

Abstract

The Gini index is the most common tool to measure inequality into two distributions. Traditionally, the Gini index and the curve of Lorenz are focused on inequalities measures in the income distribution between countries or regions. But, in the last years, several authors have shown some limitations of the Gini index. In particular, it's less sensitive to inequality at the tail of income distribution. This type of problem in the Gini index could produce two types of reactions: a new reinterpretation of the Gini index and the proposal of some alternative measures to it. In this paper, we follow the previous work using the Csiszar f-divergence to propose using the α -divergences approach to analyze the differences between the Gini index approach and these alternatives. The Gini index has been applied to the measure of resource inequalities. The AR-Gini is an area-based measure of resource inequality that estimates inequalities between neighbourhoods regarding the consumption of specific consumer goods (Druckman and Jackson, 2008). The AR-Gini could be a useful tool to monitor the distributional impacts of resource-related interventions, but this indicator presents the same overcomes as the Gini index. We can use the Gini concentration coefficient as a measure of the concentration of distribution of a random variable, especially applied to time series of data. In recent years, several studies have studied environmental spending in the European Union (EU). We focus our analysis on the distribution of this type of spending between the countries of the EU. The objective of this study is to show the differences in indexes applied to the study of the distribution of the distribution of monetary resources to environmental conservation and the extension of environmental protected areas into the countries of the European Union (EU). In our comparative study, we use the Gini index and the α -divergence measure and compare the results to get the most accurate measure of the equity of the distribution.

Keywords: Gini Index, Gini coefficient, α -divergences, AR-Gini, Environmental protected areas, Conservation policies.

JEL Codes: C43, D63, H53, Q51.



1. Introduction

Analysis of divergence between two elements is a very common problem in the scientific literature. Various divergence measures are used to show how elements are separated (Amari, 2009).

There are several examples on trying to optimize the differences between two distributions and one of those is in politics - an open problem to establish proportional methods to mitigate the unavoidable distortion arising when votes are translated to seats (Pretolani, 2013). The divergence of the distribution of votes and the distribution of seats was analysed using several measures, but, more recently, some authors evidence preference in considering the use of the Lorenz Curve and Gini Index (Mouzon et al., 2020) or the α -divergence methods (Wada & Yamahara, 2018). Another example it's the analysis of equity in the income distribution in a country. The Gini Index is a common measure of these distortions and is used to compare the situation in several countries. Using this Index is related to the Lorenz Curve (Kakwani & Podder, 1973) which depicts a graphical representation of the divergence of the distribution of two elements, for example, population and income (Rogerson, 2013).

Applications of the Gini index to environmental data can be seen in (Ivanova & Wood, 2020), where the authors consider the distribution of household carbon footprints and consumption within 27 European Union (EU) countries.

In this paper, we analyse the divergence between the distribution of environmental spending in the EU and the distribution of natural protected areas in the countries of the EU. Previous papers analyse the data divergence using an area-based measure (Druckman & Jackson, 2008) or the share of public expenses on environmental protection in total public spending in the EU (Rokicki et al., 2019). Following these papers results, now we consider the Gini Index to measure the divergence between the environmental spending of the EU countries based on the protected area.

The rest of the paper contains three sections. The next section is a review of the literature on the Gini Index and environmental spending. In the third section, we proceed to the calculation of the Gini Index and the Lorenz Curve and distortions of these values when there are small differences between distributions, in particular, tail differences. We propose the use of alternative approaches to avoid/overcome this problem. Section 4 considers the concentration of environmental spending in the EU countries. We analyse data from the period 2014-2020. Conclusions are described in the last section.

2. The Gini Index and the environmental spending

The Gini Index (Gini, 1912) is one common measure to represent income inequality and compare it between regions (Mahinpei, 2020). Although focused on income distribution, it could be used to measure the inequality of a distribution, not only in socioeconomics, but considering applications to several disciplines of science, like astrophysics, ecology, econophysics, engineering, finance, geography, medicine, biology, and transport (Siththiyot & Holasut, 2020). Interpretation of the Gini Index is very simple: the nearest to zero the more equity distribution. The maximum value is 1, and it points out that the whole welfare belongs to one agent; it represents the maximum level of inequality.



Calculation of the Gini index evolves from the original approach of Gini. The first alternative approaches: The absolute Mean Difference (Dalton, 1920) and the Relative Mean Difference Approach (Kendall & Stuart, 1958) produce the same values for the index. But the Geometric Approach (Lows, 1984) and Gini Coefficient (Dixon et al., 1987) produce different values and it could be used to improve the accuracy of the index.

The Gini Index can be derived from the Lorenz curve (Lorenz, 1905). Lorenz curve is a graphical representation of wealth distribution that plots the cumulative proportion of the population on the x-axis and the cumulative proportion of income on the y-axis, both from the lowest to the highest. The Gini Index is the ratio of the area between the area closed between the equity line and the Lorenz Curve and the total area between the equity line. Let $F(x)$ be the cumulative distribution function (cdf) of a continuous random variable X , then:

$$L(p) = \frac{\int_0^p F^{-1}(y)dy}{\int_0^1 F^{-1}(y)dy}, 0 \leq p \leq 1 \quad (1)$$

where $F^{-1}(y) = \sup\{x: F(x) \leq y\}$, is the Lorenz Curve (LC) associated with $F(x)$. Lorenz Curve satisfies the following conditions: $L(0) = 0, L(1) = 1$. The LC function $L(p)$ is an increasing and convex function for $p \in (0,1)$. Several authors consider the use of parametric models to calculate LC, the most used approaches consider the use of one parameter (Chotikapanich, 1993), two parameters (Ortega et al., 1991) and three parameters (Sarabia et al., 1999). Results indicate that the more parameters the more accuracy (Castillo et al., 1998).

The increased importance of environmental protection in people's awareness has also led to the need to count cash flows in this area (Rokicki et al., 2019). These authors show the level of public spending on environmental protection in the European Union countries. The analysed period corresponds from 2005 to 2017, and the first conclusion is that the public expenditure on environmental protection in EU countries has been systematically growing, and there was a very high concentration of this type of expenditure (Gini coefficient = 0.73) in several EU countries (Rokicki et al., 2019). Although there were changes in the value of total public expenditure for environmental protection, there were virtually no changes in the structure. The countries with the largest expenditure on environmental protection still had a dominant position (Rokicki et al., 2019).

In 2017, Ercolano and Romano published a paper where the authors analysed the national spending policies for environmental protection: "This paper has empirically examined trends in government expenditure for environmental protection in recent years at the European level, analysing whether or not European countries significantly diverge in the composition of environmental public expenditure, and identifying those countries that display greater similarities in their spending behaviour concerning environmental protection." (Ercolano & Romano, 2017). The authors use Cluster Analysis and Principal Component Analysis to create several clusters of similar countries. In line with other studies, the authors confirm that there is no convergence between environmental spending in the EU countries.

Some previous studies consider the concept and methodology for an area-based indicator of inequalities in resource use: the AR-Gini (Druckman & Jackson, 2008). The AR-Gini differs from the Gini coefficient: it's a measure of inequality in terms of the mass of resources instead of being a monetary measurement, and it's calculated on an area basis, giving a measure of inequality by



comparing the resource uses of neighbourhoods, whereas the conventional Gini compares on a household or per capita basis (Druckman & Jackson, 2008). Their study focused on the UK case and considered the period from 1996/7 to 2003/4. The paper is focused on illustrating that different consumer commodities have different inequalities that are quantified by the AR-Gini index. The results may have several implications for public spending policies.

Our paper is focused on the EU environmental spending distributed by the countries of the EU, but we consider the distribution of spending based on the protected area of the country: The bigger the protected area, the higher the level of environmental spending.

3. The Gini Index and the Lorenz Curve

The Gini index is a measure of statistical dispersion introduced by Corrado Gini (1912). Traditionally in economics, it's used as a measure of the concentration of income among the population in a country (Xu, 2004).

The Lorenz Curve (Lorenz, 1905) is a curve that defines the cumulative proportion of income and the cumulative proportion of income-receiving units (Kakwani & Podder, 1973).

The Gini Index is calculated as the area between the equity line and the Lorenz Curve (LC). However, there are several approaches to calculating this area, and we must choose high accuracy to get the right conclusions from our analysis.

3.1. The Gini Index

The first approach to the Gini Index considers the differences between the distribution of two variables. Let be two distributions: Cumulative Proportion of Population x_p and Cumulative Proportion of Income y_p . Then, we can use the following approaches to calculate the Gini Index:

1. Gini's formula: x_p presents n equal groups and μ_y is the average income: $\frac{\sum_{i=1}^n y_i}{n}$:

$$G = \frac{1}{2\mu_y n^2} \sum_{i=1}^n (n+1-2i)(y_{n-i+1} - y_i) \quad (2)$$

Mean Absolute Differences: We consider n equal groups of the population and μ_y is the mean:

$$G = \frac{\sum_i \sum_j |y_i - y_j|}{2\mu_y n^2} \quad (3)$$

Brown's Formula: We consider n equal groups of the population x_i and the income y_i :

$$G = 1 - \sum_{i=1}^{n-1} (x_{i+1} - x_i)(y_{i+1} + y_i) \quad (4)$$

But these approaches present some limitations: the limitation of the Gini index in ranking income inequality when the Lorenz curves of the two distributions cross and another limitation is whenever two or more distributions share the same value of the Gini index but distribution inequality among them could be different (Sitthiyot & Holasut, 2020).



We consider the following 2 scenarios of income distribution into a region:

Table 6: Scenario 1

District	Population	Income	Acum. Pop.	Acum. Income	Aver. Income
1	2000	1000000	0,33	0,16	500
2	2000	1000000	0,66	0,33	500
3	1000	1000000	0,83	0,5	1000
4	1000	3000000	1	1	3000

Table 7: Scenario 2

District	Population	Income	Acum. Pop.	Acum. Income	Aver. Income
1	3000	1000000	0,5	0,16	333.33
2	1000	1000000	0,66	0,33	1000
3	1000	2000000	0,83	0,66	2000
4	1000	2000000	1	1	2000

In Table 1, the first scenario, we can observe that people with low-income (≤ 500) is 4000, which represents 66,66% of the total population. But in Table 2, the second scenario, the low-income people are only 3000, which represents, 50% of the total population. In Scenario 1, we can observe that the first and second districts (the most populated) have an average income of only 500, and the fourth district has an average income of 3000. But in Scenario 2, we increase the population of District 1, which reduces the average income to only 333.3, but we move partially the income of Districts 3 and 4 to equalize them. The results are that the difference between the low-income people and the high-income people ranges from 2500 to 1666,67.

However, the two Scenarios have the same value for the Gini Index: 0.3889 (calculated using Brown's Formula (4)). We can observe that the Gini index is not fully sensitive to changes in the tails of the distribution, a conclusion that aligns with (Amari, 2009).

3.2. The Lorenz Curve

Some authors (see Rogerson, 2013) consider a new interpretation of the Gini Index from the Lorenz Curve because the Lorenz Curve is more sensitive to changes in the tails of distributions. Let $F(x|\alpha)$ be the cumulative distribution function (cdf) of a continuous random variable x , following a distribution with parameter α . Then, the Lorenz curve is:

$$L(x, \alpha) = \frac{1}{\mu} \int_0^x F^{-1}(y|\alpha) dy, \quad x \in [0,1] \quad (5)$$

where μ is the mean of F and $F^{-1} = \inf\{z: F(z|\alpha) \geq y\}$.

Using this geometrical approach the Gini Index is:

$$G = 1 - \frac{2}{\mu} \int_0^{+\infty} z F(z|\alpha) f(z|\alpha) dy = 1 - 2 \int_0^1 L(y|\alpha) dy \quad (6)$$



To improve the performance of the Lorenz Curve, we can use several distributions and increase the number of parameters to consider.

Following Castillo et al. (1998) some approaches of the Lorenz curve could be defined from different base distributions:

1. Lorenz Curve with 1 parameter. Pareto Distribution.

$$L_{CP}(x|\alpha) = 1 - (1-x)^{\frac{\alpha-1}{\alpha}}, \quad \alpha > 1 \quad (7)$$

2. Lorenz Curve with 1 parameter. Chotikapanich (1993).

$$L_1(x|\alpha) = \frac{\exp(\alpha x) - 1}{\exp(\alpha) - 1}, \quad \alpha > 0. \quad (8)$$

3. Lorenz Curve with 2 parameters. Ortega et al. (1991).

$$L_2(x|\alpha, \beta) = x^\alpha (1 - (1-x)^\beta), \quad \alpha \geq 0, \quad 0 \leq \beta \leq 1. \quad (9)$$

4. Lorenz Curve with 3 parameters. Sarabia et al. (1999).

$$L_3(x|\alpha, \beta, \gamma) = x^\alpha (1 - (1-x)^\beta)^\gamma, \quad \alpha \geq 0, \quad 0 \leq \beta \leq 1, \quad \gamma \geq 1. \quad (10)$$

The percentile method is used to estimate the parameters and fit the Lorenz curves to observed data (Castillo et al., 1998). “The results indicate that proposed methods perform as good as the currently used methods ... but in many other cases they outperform current methods ...” (Castillo et al., 1998)

Other approaches (Chotikapanich & Griffiths, 2002) consider estimating the Lorenz Curve assuming a particular distribution: Dirichlet Distribution (Minka, 2000). “Maximum likelihood estimates under the Dirichlet Distribution assumption provide better fitting Lorenz curves than nonlinear least squares and another estimation technique that has appeared in the literature” (Chotikapanich & Griffiths, 2002). An alternative approach (Chotikapanich & Griffiths, 2005) uses the Bayesian model averaging to estimate the parameters of the Lorenz Curve. A more sophisticated method of estimating generalized beta distribution parameters based on grouped data is in Kakamu & Nishimo (2019), where the authors consider the application of their method to the accuracy of the Gini coefficients estimation. A new Bayesian approach to estimating the Gini coefficient and the Lorenz curves is in Kobayashi & Kakamu (2019).

In Table 3, we compare the different values for the Gini Index calculated for the two Scenarios defined in Table 1 and Table 2. We have used the Maximum likelihood estimation (MLE) (see Chotikapanich & Griffiths, 2002). Results show that for the Gini Index estimated for a Lorenz curve with only one parameter the two Scenarios present a very similar value for the Gini Index but, increasing the number of the parameters of the estimated Lorenz Curve, produces bigger differences for the Gini Index.



Table 8: The Gini Index – Scenario 1 and Scenario 2

Lorenz Curve	Scenario 1	Scenario 2
One Parameter (Chotikapanich)	0.410503	0.413772
Two Parameters (Ortega et al.)	0.425199	0.433045
Three Parameters (Sarabia et al.)	0.408994	0.518753

For the whole three approaches, the Gini Index is worse in Scenario 2 than in Scenario 1. This fact is common to the three Lorenz Curves, and it is not shown when we use Brown’s Formula for the Gini Index. In Figure 1 to Figure 3, we depict the Lorenz Curves for the different scenarios.

Figure 3: Lorenz Curve (Chotikapanich) – Scenario 1 and Scenario 2.

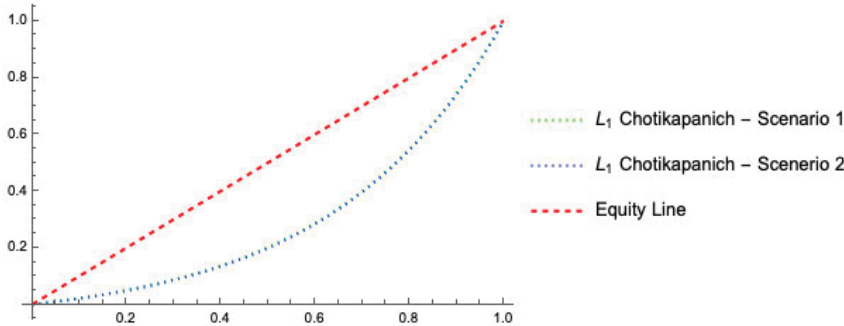


Figure 2. Lorenz Curve (Ortega et al.) – Scenario 1 and Scenario 2.

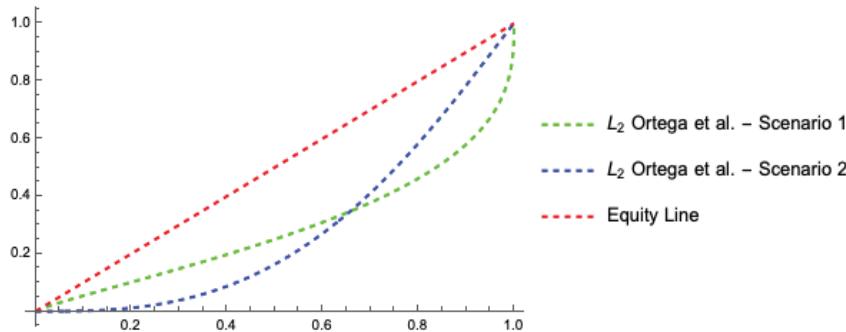
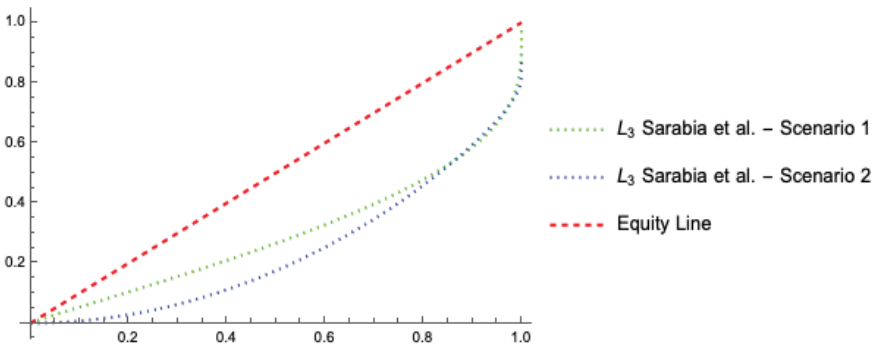


Figure 3. Lorenz Curve (Sarabai et al.) – Scenario 1 and Scenario 2.





The value of the Gini Index is sensitive to the number of parameters used on calculations of the Lorenz Curve. With only one parameter, the values for the two Scenarios are very similar, although the greater value of Scenario 2 indicates that the concentration is greater in the second scenario. When two parameters are considered, the difference is greater, and the second scenario presents a greater concentration of income. Finally, for the Lorenz Curve with 3 parameters, the difference is bigger than 0.1 in absolute value.

The Lorenz Curve is more sensitive to changes. Between Scenario 1 and Scenario 2, we have made two modifications. For low-income people we have increased the population for the first quantile, and for the high-income people we have made a redistribution of the income between the two last percentiles. This change is observed in Figure 2, Figure 1 presents the same curve and the differences in Figure 3 are only for low-income people.

4. Environmental Spending in the EU.

Environmental protection has been increasing in the EU during the last few years. The influence of people's awareness produces the need to increase cash flows on this area (Rokicki et al., 2019). The analysis of environmental protection expenditures is of strategic importance and allows the assessment of existing environmental policies. A low level of expenditure does not always mean that the country does not effectively protect its environment (Rokicki et al., 2019). Some authors associate expenditure with the level of adverse effects on nature (Rokicki et al., 2019). But in this paper, as we already mentioned, we consider the relation between the expenditure and the protected surface area.

The research period of our study concerned the years 2014-2020, and we have considered the 27 countries of the EU. The available data used were EUROSTAT data with two specific tables. Spending data were in the table National Expenditure Environmental Protection (S1) data. This table reports the environmental total spending in millions of Euros. For the surface of the protected area, we use the table Natura 2000 Protected Areas. This table reports the protected area of the EU countries in square kilometres (km²).

The research period of our study concerned to the years 2014-2020, and we have considered the 27 countries of the EU. The available data used were EUROSTAT data with two specific tables. Spending data were in the table National Expenditure Environmental Protection (S1) data. This table reports the environmental total spending in millions of Euros. For the surface of the protected area, we use the table Natura 2000 Protected Areas. This table reports the protected area of the EU countries in square kilometres (km²).

For analysing the concentration of the spending, we studied the Gini Index and the Lorenz Curve of the environmental spending, according to the total protected surface. During the analysed period the protected surface increased from 760353 km² in 2014 to 764222 km² in 2020, which represents about 0,51%. But for the same period the environmental spending increased from 238559.5 million euros to 297492.2 million euros, that is, about 24,7%. Those values point out the increasing level of environmental spending in the EU. Although the protected area stayed almost constant during these years, the spending increased by almost 25%, which clearly shows evidence of protection concerns and consequent changes in the political measures.

In Table 4 and Table 5, we present the data on the evolution of the protected area and the evolution of the environmental spending in the EU countries for the years 2014 to 2020. We can visualize this evolution in Figure 4.



The observed data of Protected Areas in the EU countries generally increased during the period 2014 to 2020. The 27 countries increase or conserve the same protected area for the analysed period. While Denmark, Hungary and Latvia kept the same area in 2017 as in 2020, Ireland and Slovenia decreased a little bit, Netherlands also decreased the area, but the worst result was in Portugal, where there was the biggest decrease in the Protected Area verified from 2019 after, while a bigger fixed Protected Area was considered in the previous 5 years.

Table 9: Natura 2000 Protected Area in EU countries (km²)

Country	2014	2015	2016	2017	2018	2019	2020
Austria (AT)	12691	12691	12691	12701	12891	12895	12895
Belgium (BE)	3887	3887	3887	3892	3891	3891	3894
Bulgaria (BG)	38222	38222	38222	38222	38222	38728	38728
Cyprus (CY)	1653	1653	1653	1653	1654	1669	1680
Czechia (CZ)	11061	11061	11148	11148	11148	11148	11149
Germany (DE)	55170	55170	55200	55200	55214	55228	55228
Denmark (DK)	3594	3594	3594	3594	3616	3594	3594
Estonia (EE)	8083	8083	8083	8106	8106	8106	8106
Greece (EL)	35747	35747	35747	35982	35982	35982	35982
Spain (ES)	137757	137757	137872	137952	138016	138111	138083
Finland (FI)	42495	42495	42495	42495	42495	42495	42498
France (FR)	69974	69974	70515	70625	70750	70875	71030
Croatia (HR)	20704	20704	20704	20704	20704	20716	20716
Hungary (HU)	19949	19949	19949	19949	19949	19949	19949
Ireland (IE)	9227	9227	9226	9230	9229	9226	9225
Italy (IT)	57172	57172	57173	57265	57265	57258	57354
Lithuania (LT)	7938	7938	8086	8086	8103	8136	8185
Luxembourg (LU)	702	702	702	702	702	702	702
Latvia (LV)	7446	7446	7446	7447	7447	7447	7446
Malta (MT)	41	41	41	41	42	42	42
Netherlands (NL)	5518	5518	5520	5522	5522	5522	5495
Poland (PL)	61165	61165	61165	61165	61156	61168	61220
Portugal (PT)	19010	19010	19010	19010	19010	18968	18968
Romania (RO)	53781	53781	54214	54214	54214	54214	54214
Sweden (SE)	55250	55250	55280	55647	55611	55611	55534
Slovenia (SI)	7674	7674	7675	7675	7672	7672	7672
Slovakia (SK)	14442	14442	14442	14633	14633	14633	14633

In general, the National Expenditure increased by almost 25% during the years 2014 to 2020. The increase was very high for the whole of the countries, but the distribution of this spending varies during the period. In 2014 the country with the lowest amount of environmental expenditure was Malta (139.4 million euros), and the highest spending was Germany (63744.5 million euros). But in 2020, the environmental expenditure of Malta increased to 307.3 million euros, (the second lowest) and then Cyprus with only 230.6 million euros occupied the lowest position. In the highest position, there was no difference in the three highest countries (Germany, France, Italy). There were some modifications of the positions of the countries during the analysed period. Although most of the



countries have increased environmental expenditure, some countries have reduced the spending for the analysed years: the Netherlands reduced environmental spending from 18358 (2014) to 15039.8 (2020) a -18.07%, and it moves from the fourth highest position to the sixth. The other countries that have reduced environmental spending are Latvia (-16.62%), Hungary (-10.85%), Finland (-21.96%), Greece (-12.32%), Denmark (-8.05%), Cyprus (-16.30%). On the positive side, the highest verified increase was in Romania (365.08%) and Malta (120.44%).

We have analysed the concentration of environmental spending in the countries of the EU according to the protected area. We based our research on the available observations on cumulative proportions of protected area (x_i) and corresponding cumulative proportions of environmental spending (y_i), obtained after ordering the countries according to increasing spending. Using these data, we were able to calculate the Gini Index and depict a Lorenz Curve for the different years.

Table 10 : National Expenditure Environmental Protection EU countries (Million €)

Country	2014	2015	2016	2017	2018	2019	2020
Austria (AT)	10211.4	10454.4	11320.5	12544.7	13454	13875.1	13071.1
Belgium (BE)	13341.7	13103.6	13286.7	14438.6	15292.4	15438.9	15985.1
Bulgaria (BG)	1096.2	982.3	662.1	764.2	711.7	924.6	1248.9
Cyprus (CY)	275.5	320.9	302.6	289.1	304.7	380.2	230.6
Czechia (CZ)	4397.3	4660	4752.6	5260.2	5775.1	5848.2	5869.9
Germany (DE)	63744.3	64748.1	67477.6	70053	74004.5	76272.8	80048.8
Denmark (DK)	5478.6	5989.2	5942	6126.4	6026	6198.8	5037.8
Estonia (EE)	525.6	588.1	498.5	555.3	635.2	620	631.5
Greece (EL)	2540.2	2368.4	2105.2	2042.6	2198.6	2278.2	2227.3
Spain (ES)	16281.4	16891.7	16985	18479.4	19494.6	19892.7	19386.9
Finland (FI)	3543.6	3746.1	3776.6	3694.3	3568.5	3388	2765.3
France (FR)	42795.2	42473.5	42651.3	44070.2	45615.6	48379.7	46726.2
Croatia (HR)	1029.4	1065.7	1102.1	1192.7	1218.5	1250.5	1005.2
Hungary (HU)	2471.6	2778.4	2111	2436.9	2123.3	2487.2	2203.5
Ireland (IE)	2179	2069.7	2032.9	2017.9	2052.3	2325.7	2302.9
Italy (IT)	26550.6	27861.9	31075.8	31887.3	32758.8	32987.1	43102.7
Lithuania (LT)	658.1	638.4	710	703.8	858.3	851.2	758.5
Luxembourg (LU)	436.5	460.8	501.3	567	624.4	617.6	815.2
Latvia (LV)	462.7	427.1	444.3	431.2	487.6	607.8	385.8
Malta (MT)	139.4	196.2	142	153.8	175.1	208.6	307.3
Netherlands (NL)	18358	18710.9	17813.8	17930.2	18968	19329.4	15039.8
Poland (PL)	7457.2	7970.1	8665.7	8815.8	9217.6	15045.4	14743.8
Portugal (PT)	2482	2442.2	2289.4	3175.6	3430.2	3595.8	3334.4
Romania (RO)	1577.4	2323.2	1386.9	1220.5	1585.3	1837	7336.2
Sweden (SE)	8180.6	8505	9089.3	9514.4	9790.2	9823.2	10004.4
Slovenia (SI)	891.1	822.8	810.9	838	917.3	1001.9	1162.3
Slovakia (SK)	1454.9	1826	1520.1	1592.9	1536.2	1671	1761.3

Using Brown's approach to the Gini Index it is possible to calculate the value of the Gini Index as:

$$GI = 1 - \sum_{i=1}^{n-1} (y_{i+1} + y_i)(x_{i+1} - x_i) \quad (11)$$

The evolution of the data (see Table 6) shows that, for the whole period, concentration of the spending was decreasing. The Gini Index moves from 0.45 (2014) to 0.42 (2020). Although there was a bit increase in 2016, the other years show that the trend is diminishing. There is a better distribution of spending according to the protected area surface in 2020 than in 2019, although the value of the Gini index is high (>0.4) which points out that the distribution of spending is far from equity.

As we have commented, Brown's formula for the Gini Index presents some problems of insensitive changes in the tail of the distribution. So, we considered another approach to the Gini Index by using an estimation of the Lorenz Curve with two parameters $y = L(x; \alpha, \delta) = x^\alpha(1 - (1 - x)^\delta)$ with $\alpha \geq 0, 0 \leq \delta \leq 1$. (Ortega et al., 1991).

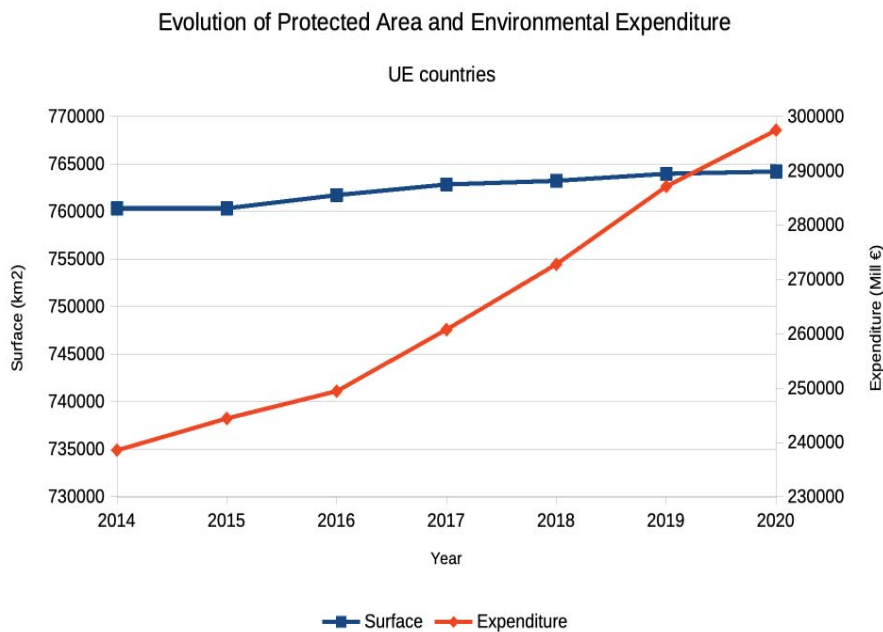
The spending shares can be defined as $q_i = y_i - y_{i-1}$, and supposing they are random variables with means $E(q_i) = L(x_i, \alpha, \delta) - L(x_{i-1}, \alpha, \delta)$, following a Dirichlet distribution (Chotikapanich & Griffiths, 2002). Thus, the probability density function for q can be written as:

$$f(q | \lambda, \alpha, \delta) = \Gamma(\lambda) \prod_{i=1}^{28} \frac{q_i^{\lambda(L(x_i; \alpha, \delta) - L(x_{i-1}; \alpha, \delta)) - 1}}{\Gamma(\lambda(L(x_i; \alpha, \delta) - L(x_{i-1}; \alpha, \delta)))} \tag{12}$$

There is then the need to estimate the parameters: λ, α, δ , where α, δ are the Lorenz Curve parameters and λ acts as an inverse variance parameter. The larger the value, the better the fit of the Lorenz curve to the data. Taking into account the Maximum Likelihood Estimation, the values of parameters can be obtained by maximizing the log-likelihood function:

$$\text{Log}[f(q | \lambda, \alpha, \delta)] = \text{Log } \Gamma(\lambda) + \sum_{i=1}^{28} (\lambda(L(x_i; \alpha, \delta) - L(x_{i-1}; \alpha, \delta)) - 1) \text{Log } q_i - \sum_{i=1}^{28} \text{Log } \Gamma(\lambda(L(x_i; \alpha, \delta) - L(x_{i-1}; \alpha, \delta))) \tag{13}$$

Figure 4: Evolution Spending/Surface EU (2014-2017)





The Gini Index is defined as

$$GI = 1 - 2 \int_0^1 L(x; \alpha, \delta) dx \tag{14}$$

Table 6: The Gini Index

Gini Index	2014	2015	2016	2017	2018	2019	2020
Brown's Formula	0.4589	0.4515	0.4676	0.4388	0.4381	0.4235	0.4212
L ₂ (Ortega et al., 1991)	0.4936	0.4886	0.4747	0.4769	0.4782	0.4759	0.4324

Figure 5: Lorenz Curve Distribution Environmental Expenditure EU (2014-2017)

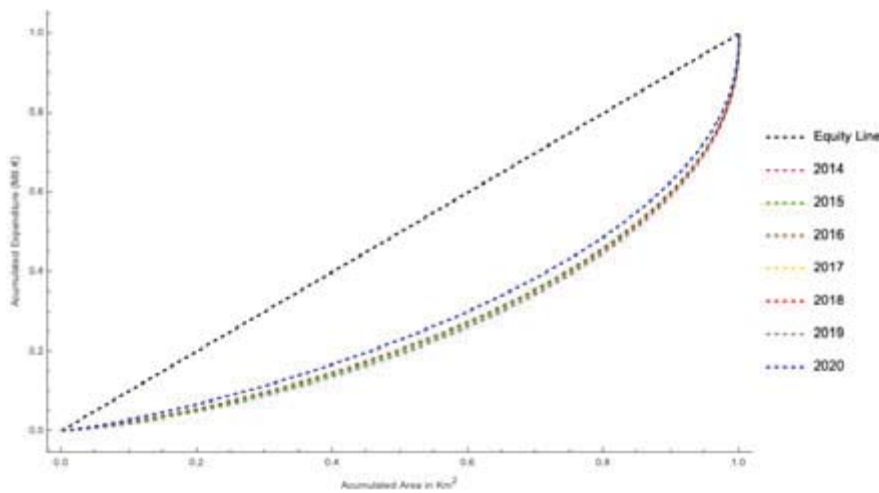
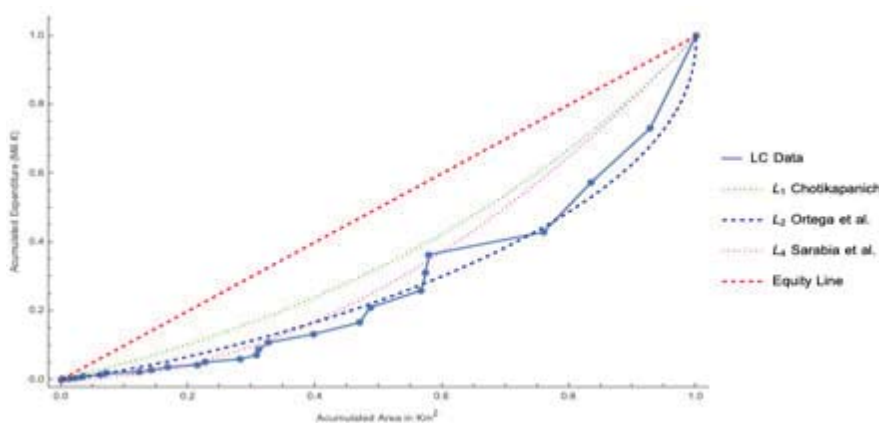


Figure 6: Lorenz Curves using different approaches (2020)



As we can observe (see Table 6) the concentration of spending was diminishing along the years 2014 to 2020, moving from 0.49 (2014) to 0.43 (2020). The value of concentration is high, although the reduction of it into the analysed period is about 0.06 points. Lorenz Curve for several years is depicted in Figure 5.

There are other approaches to estimate the Lorenz Curve using different numbers of parameters. In Figure 6, we have represented the estimation of different approaches to the Lorenz



Curve, and we have compared it with the Equity Line and with the Real Data for the year 2020. Graphical results show that L_2 is the most accurate curve to real data.

5. Conclusions

The Gini Index is the most used index to analyse the divergence between two data series. Usually, the Gini Index is used to analyse the distribution of income between populations in a country. A comparative study between countries was performed, based on the Gini Index values.

However, the optimal version for the Gini Index is an open problem, let's call it OGIV (Optimal Gini Index Version), and it is our aim to continue our research looking for it. The original Gini Index formula produces values that are not sensitive to small changes, especially on the tails of the distribution. Some authors consider that a more accurate value is reached when we use the Lorenz Curve and consider one or several parameters to define the most accurate curve. Then, the Gini Index is calculated by integrating the area between the Lorenz Curve and the equity line.

In this paper, we have used the Gini Index to measure the concentration of the national environmental expenditure based on the national protected area for the 27 countries of the EU, during the years 2014-2020. We have used the L_2 Lorenz Curve approach (Ortega et al., 1991) to calculate the Gini Index of the different years. Parameters of the model were calculated using Maximum Likelihood Estimation and Dirichlet Distribution.

The data show that the evolution of national expenditures, in general, during these years increased by about 25%, although the protected area remains very stable with a general little increase (0.5%). The distribution of the spending reflects some changes between the different countries' policies at which may possibly be one of the reasons for some variations in the order of the countries year by year.

The Gini Index diminished from 0.49 to 0.42, that is, a fall in the concentration of the spending. Although the concentration is very high, it seems that the evolution of the expenditure tends to have a more equity distribution between the countries based on the protected area.

We have considered an alternative approach to estimate the Lorenz Curve, only for the year 2020. Finally, we observed that the most accurate approach to the Real Data distribution was given by using the L_2 curve.

Acknowledgement

This work is partially financed by national funds through FCT – Fundação para a Ciência e a Tecnologia under the project UIDB/00006/2020.

DOI: 10.54499/UIDB/00006/2020 (<https://doi.org/10.54499/UIDB/00006/2020>).

References

1. Amari, Shun-Ichi. (2009). α -Divergence Is Unique, Belonging to Both f -Divergence and Bregman Divergence Classes. *IEEE Transactions on Information Theory*: 55 (11), 4925-31.
2. Castillo, Enrique, Ali S. Hadi, and José María Sarabia. (1998). A Method for Estimating Lorenz Curves. *Communications in Statistics - Theory and Methods*: 27 (8), 2037-63.



3. Chotikapanich, Duangkamon, and William E. Griffiths. (2002). Estimating Lorenz Curves Using a Dirichlet Distribution. *Journal of Business and Economics Statistics*: 20 (2), 290-95.
4. Cohtikapanich, Duangkamon, and William E. Griffiths. (2005). Averaging Lorenz Curves. *Journal of Economic Inequality*: 3, 1-19.
5. Dalton, Hugh. (1920). The Measurement of the Inequality of Incomes. *The Economic Journal*: 30 (119), 348-61.
6. Dellmuth, Lisa M., and Adam W. Chalmers. (2018). All Spending Is Not Equal: European Union Public Spending, Policy Feedback and Citizens' Support for the EU. *European Journal of Political Research*: 57, 3-23.
7. Dixon, Philip M., Jacob Weiner, Thomas Mitchell-Olds, and Robert Woodley. (1987). Bootstrapping the Gini Coefficient of Inequality. *Ecology*: 68 (5), 1548-51.
8. Druckman, A., and T. Jackson. (2008). Measuring Resource Inequalities: The Concepts and Methodology for an Area-Based Gini Coefficient. *Ecological Economics*: 65, 242-52.
9. Ercolano, Salvatore, and Oriana Romano. (2018). Spending for the Environment: General Government Expenditure Trends in Europe. *Social Indicators Research*: 138, 1145-69.
10. Gini, C. (1912). Variabilità e Mutuabilità. *Contributo allo Studio delle Distribuzioni e delle Relazioni Statistiche*. Bologna: C. Cuppini.
11. Gini, C. (1921). Measurement of Inequality of Incomes. *The Economic Journal*: 31 (121), 124-26.
12. Ivanova, Diana, and Richard Wood. (2020). The Unequal Distribution of Household Carbon Footprints in Europe and Its Link to Sustainability. *Global Sustainability*: 3 (18), 1-12.
13. Kakamu, Kazuhiko, and Haruhisa Nishino. (2019). Bayesian Estimation of Beta-Type Distribution Parameters Based on Grouped Data. *Computational Economics*: 54, 625-45.
14. Kakwani, N. C., and N. Podder. (1973). On the Estimation of Lorenz Curves Form Grouped Observations. *International Journal Economic Review*: 14 (2), 278-92.
15. Kendall, M.G. and Stuart, A. (1958). *The Advance Theory of Statistics*. 1. London: Griffin.
16. Kobayashi, Genya, and Kazuhiko Kakamu. (2019). Approximate Bayesian Computation for Lorenz Curves from Grouped Data. *Computational Statistics*: 34, 253-79.
17. Lorenz, M. O. (1905). *Methods of Measuring the Concentration of Wealth*. Publications of the American Statistical Association: 9(70), 209-219.
18. Lows, Raymond L. (1984). Measurement of Inequality: The Gini Coefficient and School Finance Studies. *Journal of Education Finance*: 10 (1), 83-94.
19. Mahinpei, Romina. (2021). A Comparative Analysis of the Gini Index." *STEM Fellowship Journal*: 6 (1), 34-52.
20. Minka, Thomas P. (2000). "Estimating a Dirichlet Distribution." <https://tminka.github.io/papers/dirichlet/minka-dirichlet.pdf>
21. Mouzon, Olivier de, Thibault Laurent, and Michel Le Breton. (2020). One Man, One Vote Part 2: Measurement of Malapportionment and Disproportionality and the Lorenz Curve. *Université de Toulouse Capitole*.
22. Pretolani, Daniele. (2014). Apportionments with Minimum Gini Index of Disproportionality: A Quadratic Knapsack Approach. *Annals of Operations Research*: 215, 257-67.
23. Rogerson, Peter A. (2013). The Gini Coefficient of Inequality: A New Interpretation. *Letters in Spatial and Resource Sciences*: 6 (3), 109-20.
24. Rokicki, Tomasz, Luiza Ochnio, Grzegorz Koszela, Agata Zak, Edyta Karolina Szczepaniuk, Hubert Szczepaniuk, Konrad Michalski, and Aleksandra Perkowska. (2019). Public Expenditure on Environmental Protection in the European Union Countries. *Middle Pomeranian Scientific Society of the Environment Protection*: 21, 364-77.



25. Sarabia, José María, Enrique Castillo, and Daniel J. Slottje. (1999). An Ordered Family of Lorenz Curves. *Journal of Econometrics*: 91, 43-60.
26. Sitthiyot, Thititthep, and Kanyarat Holasut. (2020). A Simple Method for Measuring Inequality. *Palgrave Communications*: 6 (1), 1-9.
27. Wada, Junichiro. (2016). Apportionment Behind the Veil of Uncertainty. *Japanese Economic Review*: 67 (3), 348-60.
28. Wada, Junichiro and Yuta Yamahara (2018). Studying malapportionment using α -divergence. *Mathematical Social Sciences*: 93, 77-89.
29. Xu, Kuan. (2004). How Has the Literature on Gini's Index Evolved in the Past 80 Years? Dalhousie University, Economics Working Paper.