METRON
https://doi.org/10.1007/s40300-024-00274-7

On the decomposition of the extensions of the Gini index that are based on the 'metallic' sequences of number theory

Elena Bárcena-Martin[1] · **Jacques Silber**[2]

Received: 20 September 2023 / Accepted: 22 July 2024

Abstract

The present paper extends the work of Subramanian (Econ Bull 41(4):2309–2319, 2021), as well as that of Creedy and Subramanian (Exploring a new class of inequality measures and associated value judgements: Gini and Fibonacci-type sequences, 2022), by showing how this new extension of the Gini index may be decomposed by income sources, income classes and population subgroups. It also gives an empirical illustration applying this extension of the Gini index to the analysis of the inequality in expenditure on social protection among European countries in 2018. In the decomposition by benefit function it appears that the housing function consistently exhibits the smallest contribution, due to its low expenditure weight, while the old-age function has the lowest inequality but the highest contribution, due to its substantial expenditure weight. The family/children function's contribution increases when a higher weight is given to lower expenditures, highlighting a pronounced inequality at lower expenditure levels. Conversely, the old-age function shows higher inequality among countries with higher expenditures. In the breakdown by welfare systems, we observe greater inequality between systems than within them.

European Union · Fibonacci · Gini · inequality decomposition · social protection

---

This paper was written in memory of the late Giovanni Maria Giorgi who was, without any doubt, the greatest expert on the literature related to the Gini index. Elena Barcena and Giovani Giorgi first crossed paths within the online academic research networks. Giovani reached out to Elena to apprise her of his research endeavors concerning the Gini index and its historical context. From that initial contact, they have maintained an ongoing and amicable scholarly exchange, regularly sharing recent publications pertaining to the Gini index and related inequality measures, including those of Bonferroni and De Vergotini. Whenever Jacques Silber wrote a paper somewhat related to the Gini index, he would always make sure to send a copy of the paper to Giorgi. Though not close friends, Jacques and Giovanni Maria met several times in academic conferences and used to send each other papers related to the Gini index. Giovanni Maria was two weeks younger than Jacques and Giovanni Maria's untimely death leaves a void to those working on inequality, in particular on the Gini index.

✉ Jacques Silber
jsilber_2000@yahoo.com

Elena Bárcena-Martin
barcenae@uma.es

[1] Facultad de Ciencias Económicas y Empresariales, Universidad de Málaga, Malaga, Spain

[2] Department of Economics, Bar-Ilan University, 52900 Ramat-Gan, Israel

1 Introduction

“Some of the many reasons for the success and the relevance of the Gini index are its simplicity, certain interesting properties and useful decomposition possibilities…” [6]. There are in fact several ways of formulating the Gini index as “it can be expressed as a ratio of two regions defined by a 45-degree line and a Lorenz curve in a unit box, or a function of Gini’s mean difference, or a covariance between incomes and their ranks, or a matrix form of a special kind” [11]. There have also been extensions of the Gini index, such as those proposed by Donaldson and Weymark [3] or Yitzhaki [12], which involve a single-parameter extension of the Gini index. More precisely, while the Gini coefficient is derived from a weighting of incomes based on the Borda rank-order system, the generalizations of the Gini index raise the income-weights to higher powers. This leads to an extension of the Gini index which becomes ‘distribution sensitive’.

More recently, in a very original paper, Subramanian [10] introduced an alternative generalization of the Gini index based on the notion of “metallic” sequences of number theory.

The present paper extends this work of Subramanian [10], as well as that of Creedy and Subramanian [1], by showing how this new extension of the Gini index may be decomposed by income sources, income classes and population subgroups. Section 2 summarizes the contributions of Subramanian [10] and Subramanian and Creedy [2]. The focus of Sect. 3 is on the decomposition by income sources of this new extension of the Gini index while Sects. 4 and 5 present the decomposition of this extended Gini index by income classes and population subgroups. Section 6 presents then a simple illustration while Sect. 7 gives an empirical illustration applying this extension of the Gini index to the analysis of the inequality in expenditure on social protection among European countries. Section 8 finally concludes this paper.

2 Subramanian’s single-parameter generalization of Gini based on the ‘metallic’ sequences of number theory

Extensions of the Gini index have appeared previously in the literature. For example, the so-called ‘S-Gini’ parametrization amounts to using a scalar as an exponent on Gini’s income-weight. Subramanian [10] introduced an alternative single-parameter generalization in which income-weights are derived from Fibonacci-like sequences of numbers. The Gini coefficient turns then out to be a special case of the resulting series of indices. Subramanian [10] introduced also the ‘Fibonacci’ index which is shown to be a transfer-sensitive extension of the Gini index.

Let us first remind the reader of what a ‘metallic’ sequence of numbers refers to.

Given the set of the first n natural numbers $M_n = \{1, 2, \ldots, i, \ldots, n\}$ define for all $k = 0, 1, 2\ldots$ the mapping f_k on M_n such that $f_k(1) = 1$, $f_k(2) = Max\{f_k(1), kf_k(1)\}$ and for all $i \in M_n$, except $i = 1, 2$,

$$f_k(i) = kf_k(i-1) + f_k(i-2) \tag{1}$$

It is then easy to check that when $k = 0$, we will obtain for example for $i = 1, 2, \ldots, 7, 8$, the sequence of numbers 1,1,…,1, 1.

When $k = 1$, we will obtain for example for $i = 1, 2, \ldots, 7, 8$, the sequence of numbers 1, 1, 2, 3, 5, 8, 13, 21 which is known as the Fibonacci sequence of numbers.

When $k = 2$, we will obtain for example for $i = 1, 2, \ldots, 7, 8$, the sequence of numbers 1, 2, 5, 12, 29, 70, 169, 408, which is known as the Pell sequence of numbers.

Following Creedy and Subramanian [1] define now $S_k(i)$ and S_k as

$$S_k(i) = \sum_{j=i}^{n} f_k(j) \tag{2}$$

$$S_k = \sum_{i=1}^{n} S_k(i) \tag{3}$$

Note that when $k = 0$, the sequence of numbers corresponds to the Borda weights.[1]

Given an ordered vector x of n incomes x_i with $x_1 \le x_2 \le \ldots \le x_i \le \ldots \le x_n$, define a welfare function $W_k(x)$ as

$$W^k(x) = \sum_{i=1}^{n} S_k(i) x_i \tag{4}$$

As shown by Creedy and Subramanian [1] the "equally distributed equivalent level of income" $x_{E,k}$ corresponding to this welfare function is then defined as

$$x_{E,k} = \sum_{i=1}^{n} \left[\frac{S_k(i)}{S_k} \right] x_i \tag{5}$$

Using (5) we may then derive an extended Gini index G_k as

$$\begin{aligned} G_k &= 1 - \frac{x_{E,k}}{\overline{x}} = 1 - \sum_{i=1}^{n} \left[\frac{S_k\,(i)}{S_k} \right] \left(\frac{x_i}{\overline{x}} \right) = \sum_{i=1}^{n} w_k(i) \left[1 - \left(\frac{x_i}{\overline{x}} \right) \right] \\ &= \sum_{i=1}^{n} w_k(i) \left(\frac{\overline{x} - x_i}{\overline{x}} \right) = \sum_{i=1}^{n} w_k(i) d_i \end{aligned} \tag{6}$$

where $w_k(i) = \left(\frac{S_k(i)}{S_k} \right)$, $d_i = \left(\frac{\overline{x} - x_i}{\overline{x}} \right)$ and $\overline{x}$ is equal to the arithmetic mean of the incomes x_i. Note that greater values of k provide more weight to the lowest values of the variable, so more importance is given to the left tail of the distribution.

Note also that when $k = 0$ the index G_0 will be simply related to the traditional Gini index as shown in Appendix A. The traditional Gini index $I_G = \frac{1}{2n^2\mu} \sum_{j=1}^{n} \sum_{i=1}^{n} \left| x_j - x_i \right|$ is expressed by Creedy and Subramanian [1] as $I_G = \left(\frac{n+1}{n} \right) - \left\{ \left[\left(\frac{2}{n^2} \right) \right] \left[\sum_{i=1}^{n} (n + 1 - i) \left(\frac{x_i}{\overline{x}} \right) \right] \right\}$

---

[1] Borda weights refer to a method in voting theory named after Jean-Charles de Borda, an 18th-century French mathematician. In this method, each voter ranks the candidates in order of preference, and candidates receive points related to their reverse rank positions. The top-ranked candidate receives a certain number of points (typically equal to the number of candidates minus one), the second-ranked candidate receives one fewer point, and so on until the last-ranked candidate receives zero points. Borda weights assign numerical values to each rank, with higher ranks receiving higher weights.

3 On the decomposition by income sources of this extension of the Gini index

Define now a (1 by n) row-vector f_k as

$$f_k = \{f_k(1), \ldots, f_k(i), \ldots, f_k(n)\} \tag{7}$$

and a (n by 1) column vector x whose transpose x' is defined as

$$x' = \{x_1, \ldots, x_i, \ldots, x_n\} \tag{8}$$

Finally define a (n by n) square matrix A whose typical element a_{ij} is equal to 1 if $i \geq j$, to 0 otherwise.

It is then easy to check that $x_{E,k}$ may be expressed as

$$x_{E,k} = (f_k Ax)/S_k \tag{9}$$

while the index G_k may be expressed as

$$G_k = 1 - \left(\frac{x_{E,k}}{\overline{x}}\right) = 1 - \left(\frac{f_k Ax}{\overline{x} S_k}\right) \tag{10}$$

Let us now define a row vector p_k as

$$p_k = \left\{\left(\frac{f_k(1)}{S_k}\right), \ldots, \left(\frac{f_k(i)}{S_k}\right), \ldots, \left(\frac{f_k(n)}{S_k}\right)\right\} \tag{11}$$

a column vector r whose transpose r' is expressed as

$$r' = \left\{\left(\frac{x_1}{\overline{x}}\right), \ldots, \left(\frac{x_i}{\overline{x}}\right), \ldots, \left(\frac{x_n}{\overline{x}}\right)\right\} \tag{12}$$

and a column vector e whose transpose e' is written as

$$e' = \{1, \ldots, 1, \ldots, 1\} \tag{13}$$

Using (10) the index G_k, that is a unit vector, may then be expressed as

$$G_k = p_k Ae - p_k Ar = p_k Ad \tag{14}$$

where

$$d = e - r = \left\{\left(1 - \frac{x_1}{\overline{x}}\right), \ldots, \left(1 - \frac{x_i}{\overline{x}}\right), \ldots, \left(1 - \frac{x_n}{\overline{x}}\right)\right\} = \{d_1, \ldots, d_i, \ldots, d_n\} \tag{15}$$

Call now x_{im} the income that individual i receives from income source m. Using (14) we can then define the index $G_{k,m}$ of income source m as

$$G_{k,m} = p_k Ad_m \tag{16}$$

where

$$d_m = e - r_m \tag{17}$$

In (17) r_m is a n by 1 column vector whose transpose r'_m is defined as

$$r'_m = \left\{\left(\frac{x_{1m}}{\overline{x}_m}\right), \ldots, \left(\frac{x_{im}}{\overline{x}_m}\right), \ldots, \left(\frac{x_{nm}}{\overline{x}_m}\right)\right\} \tag{18}$$

where x_{im} is the income that individual i receives from income source m and $\overline{x}_m$ is the average income from source m.

Let x^m be a column vector of the incomes x_{im}, ranked by increasing values and $\overline{x}^m$ is the arithmetic mean of these incomes x_{im}. Note that in general the ranking of the individuals by increasing values of the total incomes x_i will be different from the ranking of the individuals by increasing values of the incomes x_{im} they receive from income source m.

Combining expressions (14) to (17) we derive

$$G_k = p_k Ad = p_k A[e - r] = p_k A\left[(e - \left[\left(\frac{x_1}{\overline{x}}\right), \ldots, \left(\frac{x_i}{\overline{x}}\right), \ldots, \left(\frac{x_n}{\overline{x}}\right)\right]'\right] \tag{19}$$

where the superscript after the square bracket indicates that the square bracket refers to a column vector.

Let $x_i = \sum_{m=1}^{M} x_{im}$ for $i = 1 \ldots n$, where M is the number of income sources, and $\overline{x}_m = \frac{\sum_{i=1}^{n} x_{im}}{n}$. Expression (19) may be rewritten as

$$\begin{aligned} G_k &= p_k A\left[(e - \left[\left(\frac{\sum\limits_{m=1}^{M} x_{1m}}{\overline{x}_m}\left(\frac{\overline{x}_m}{\overline{x}}\right)\right), \ldots, \left(\frac{\sum\limits_{m=1}^{M} x_{im}}{\overline{x}_m}\left(\frac{\overline{x}_m}{\overline{x}}\right)\right), \ldots, \left(\frac{\sum\limits_{m=1}^{M} x_{nm}}{\overline{x}_m}\left(\frac{\overline{x}_m}{\overline{x}}\right)\right)\right]'\right] = G_k \\ &= p_k Ae - p_k A\left[\frac{\sum\limits_{m=1}^{M} x_{im}}{\overline{x}_m}\left(\frac{\overline{x}_m}{\overline{x}}\right)\right]' \end{aligned} \tag{20}$$

where $\left[\frac{\sum_{m=1}^{M} x_{im}}{\overline{x}_m}\left(\frac{\overline{x}_m}{\overline{x}}\right)\right]'$ refers to a column vector whose typical element is equal to $\frac{\sum_{m=1}^{M} x_{im}}{\overline{x}_m}\left(\frac{\overline{x}_m}{\overline{x}}\right)$.

Expression (20) may be also written as

$$G_k = p_k Ae - \left[\sum_{m=1}^{M}\left(\frac{\overline{x}_m}{\overline{x}}\right)\left[p_k A\left(\frac{x_{im}}{\overline{x}_m}\right)\right]'\right]$$

$$G_k = p_k Ae + \sum_{m=1}^{M}\left(\frac{\overline{x}_m}{\overline{x}}\right)\left[p_k Ae - p_k A\left(\frac{x_{im}}{\overline{x}_m}\right)\right] - p_k Ae \tag{21}$$

since $\sum_{m=1}^{M}\left(\frac{\overline{x}_m}{\overline{x}}\right)(p_k Ae) = p_k Ae$

We therefore end up with

$$G_k = \sum_{m=1}^{M}\left(\frac{\overline{x}_m}{\overline{x}}\right)\left[p_k Ae - p_k A\left(\frac{x_{im}}{\overline{x}_m}\right)\right] = \sum_{m=1}^{M}\left(\frac{\overline{x}_m}{\overline{x}}\right) PG_{k,m} \tag{22}$$

Note that in defining the index $PG_{k,m}$ in expression (22) the ordering of the expressions $\left(\frac{x_{im}}{\overline{x}_m}\right)$ in the column vector $p_k A\left(\frac{x_{im}}{\overline{x}_m}\right)$ is that of the total incomes x_i and not that of the incomes x_{im} from source m.

Let us however call $\widetilde{x}_{im}$ the incomes from source ranked by increasing values of the income source m itself and not by increasing values of the total income. In such a case we can define the index $G_{k,m}$ from source m as

$$G_{k,m} = p_k A e - p_k A \left(\frac{\widetilde{x}_{im}}{\overline{x}_m} \right) \tag{23}$$

We can then define a correlation coefficient $GC_{k,m}$ as

$$GC_{k,m} = \frac{PG_{k,m}}{G_{k,m}} \tag{24}$$

Combining (22) and (24) we end up with

$$G_k = \sum_{m=1}^{M} \left(\frac{\overline{x}_m}{\overline{x}} \right) G_{k,m} GC_{k,m} \tag{25}$$

This result is very similar to the traditional decomposition of the Gini index by income sources (see, for example, [4], or [9].

The index G_k is therefore equal to the weighted average of the products of the indices $G_{k,m}$ and of the correlation coefficients $GC_{k,m}$, the weights being equal to the shares of the income sources in the total income of the individuals.

4 On the decomposition by income classes

Assume we decompose the set of the total incomes x_i into a number of adjacent income classes.

Combining (14) and (15) we may rewrite expression (14) as

$$\begin{aligned} G_k &= p_k A d = p_k A \left[\frac{\overline{x} - x_i}{\overline{x}} \right] = p_k A \left[\frac{(\overline{x} - \overline{x}_m) + (\overline{x}_m - x_i)}{\overline{x}} \right] \\ &= p_k A \left[\frac{(\overline{x} - \overline{x}_m)}{\overline{x}} \right] + p_k A \left[\frac{(\overline{x}_m - x_i)}{\overline{x}} \right] \end{aligned} \tag{26}$$

Note that in (26) all the expressions in square brackets refer to a column vector of the elements which appear in these square brackets.

The first element on the R.H.S. of (26) is then clearly equal to a between groups inequality index since in the column vector of the expressions $(\overline{x} - \overline{x}_m)$ the elements will be ranked by non-increasing values as is the case for the column vector of the expressions $(\overline{x} - x_i)$.

The second expression on the R.H.S. of (26) reminds us of the way a within groups Pseudo-Gini index is defined. Recall that the Pseudo-Gini index [4] or the concentration ratio in [8] is the Gini index computed over a distribution in which the variable is not ranked by non-decreasing values. In the expressions $\left(\frac{\overline{x} - x_i}{\overline{x}} \right)$, $\overline{x}$ is the mean of x_i, given that here we ignore the group to which individual i belongs. Similarly, in the expressions $\left(\frac{\overline{x} - \overline{x}_m}{\overline{x}} \right)$, $\overline{x}$ is the mean of $\overline{x}_m$. Finally, in the expressions $\left(\frac{\overline{x}_m - x_i}{\overline{x}} \right)$, $\overline{x}_m$ can be considered as the mean of x_i once we assume that we know to which group individual i belongs.

Note however that in this second expression on the R.H.S. of (26) the elements of the column vector of the expressions $\left(\frac{\overline{x}_m - x_i}{\overline{x}} \right)$ are not ranked by non-decreasing values.

5 A decomposition by population subgroups

Assume now we have population subgroups whose income distributions overlap. Let x refer to the column vector of the incomes ranked by non-decreasing values. Let then y refer to the column vector of these same incomes but ranked first by non-decreasing values of the average incomes of the population subgroups, and then within each population subgroups by non-decreasing values of the incomes of the individuals who belong to the corresponding population subgroup. We can then use (26) to decompose the Gini-type index into between and within population inequality components. In other words, we will write that

$$p_k A\left[\frac{\overline{y}-y_i}{\overline{y}}\right]=p_k A\left[\frac{\left(\overline{y}-\overline{y}_m\right)+\left(\overline{y}_m-y_i\right)}{\overline{y}}\right]=p_k A\left[\frac{\left(\overline{y}-\overline{y}_m\right)}{\overline{y}}\right]+p_k A\left[\frac{\left(\overline{y}_m-y_i\right)}{\overline{y}}\right] \tag{27}$$

with evidently $\overline{y}=\overline{x}$ and $\overline{y}_m=\overline{x}_m\forall m$.

The Gini-type index should however be expressed as

$$\begin{aligned} G_k=p_k A\left[\frac{\overline{x}-x_i}{\overline{x}}\right]&=\left[p_k A\left[\frac{\overline{y}-y_i}{\overline{y}}\right]\right]+\left[\left(p_k A\left[\frac{\overline{x}-x_i}{\overline{x}}\right]\right)-\left(p_k A\left[\frac{\overline{y}-y_i}{\overline{y}}\right]\right)\right]\\ &=p_k A\left[\frac{\overline{y}-y_i}{\overline{y}}\right]+p_k A\left[\left(\frac{y_i-x_i}{\overline{x}}\right)\right] \end{aligned} \tag{28}$$

Combining (27) and (28) we then get

$$G_k=p_k A\left[\frac{\overline{x}-x_i}{\overline{x}}\right]=p_k A\left[\frac{\left(\overline{y}-\overline{y}_m\right)}{\overline{y}}\right]+p_k A\left[\frac{\left(\overline{y}_m-y_i\right)}{\overline{y}}\right]+p_k A\left[\left(\frac{y_i-x_i}{\overline{x}}\right)\right] \tag{29}$$

The two first terms on the R.H.S. of (29) remind us of the between groups Gini-type index and of the within groups Pseudo-Gini index while the third term is an interaction term representing the overlap between the income distributions of the different population groups.

6 A simple empirical illustration

Let the distribution of incomes x be as follows: 1, 2, 7, 10, 30.

With $k=0$, as mentioned previously, $f_k(i)=1\,\forall i$. Since $S_k(i)$ is defined as $S_k(i)=\sum_{j=i}^{n} f_k(j)$, we derive that

$$S_k(1)=5; S_k(2)=4; S_k(3)=3; S_k(4)=2; S_k(5)=1$$

so that $S_k= 5+4+3+2+1=15$.

Using the data on the income distribution given previously, we derive that

$$W_k(x)=\sum_{i=1}^{5} S_k(i)x_i=5(1)+4(2)+3(7)+2(10)+1(30)=5+8+21+20+30=84$$

$$\begin{aligned} x_{Ek}&=\sum_{i=1}^{5}\left[\frac{S_k(i)}{S_k}\right]x_i=\left(\frac{5}{15}\right)1+\left(\frac{4}{15}\right)2+\left(\frac{3}{15}\right)7+\left(\frac{2}{15}\right)10+\left(\frac{1}{15}\right)30\\ &=\frac{5+8+21+20+30}{15}=\frac{84}{15}=5.6 \end{aligned}$$

The average income $\overline{x}$ is equal to $(1 + 2 + 7 + 10 + 30)/5 = 10$ so that the inequality index I_B we obtain is equal to $\left(1 - \frac{5.6}{10}\right) = 0.44$, while it can be shown that the Gini index is equal to 0.528. Creedy and Subramanian had indeed stressed that when the number of observations is small, there will be a difference between the index I_B and the Gini index.

With $k = 1$, Creedy and Subramanian [1] have shown that $f_k(1) = 1$; $f_k(2) = 1$; $f_k(3) = 2$; $f_k(4) = 3$; $f_k(5) = 5$. Since $S_k(i)$ is defined as $S_k(i) = \sum_{j=i}^{n} f_k(j)$, we derive that

$$S_k(1) = 12; S_k(2) = 11; S_k(3) = 10; S_k(4) = 8; S_k(5) = 5$$

so that $S_k = 12 + 11 + 10 + 8 + 5 = 46$.

Using the data on the income distribution given previously, we derive that

$$W_k\,(x) = \sum_{i=1}^{5} S_k\,(i)\,x_i = 12\,(1) + 11\,(2) + 10\,(7) + 8\,(10) + 5\,(30) = 12 + 22 + 70 + 80 + 150 = 334$$

$$\begin{aligned} x_{Ek} &= \sum_{i=1}^{5}\left[\frac{S_k(i)}{S_k}\right]x_i = \left(\frac{12}{46}\right)1 + \left(\frac{11}{46}\right)2 + \left(\frac{10}{46}\right)7 + \left(\frac{8}{46}\right)10 + \left(\frac{5}{46}\right)30 \\ &= \frac{12 + 22 + 70 + 80 + 150}{46} = \frac{334}{46} = 7.2609 \end{aligned}$$

The average income $\overline{x}$ is equal to $(1 + 2 + 7 + 10 + 30)/5 = 10$
so that the inequality index I_B we obtain is equal to $\left(1 - \frac{7.2609}{10}\right) = 0.2739$

With $k = 2$, Creedy and Subramanian [1] have shown that $f_k(1) = 1$; $f_k(2) = 2$; $f_k(3) = 5$; $f_k(4) = 12$; $f_k(5) = 29$. Since $S_k(i)$ is defined as $S_k(i) = \sum_{j=i}^{n} f_k(j)$, we derive that

$$S_k(1) = 49; S_k(2) = 48; S_k(3) = 46; S_k(4) = 41; S_k(5) = 29$$

so that $S_k = 49 + 48 + 46 + 41 + 29 = 213$.

Using the data on the income distribution given previously, we derive that

$$\begin{aligned} W_k\,(x) = \sum_{i=1}^{5} S_k\,(i)\,x_i &= 49\,(1) + 48\,(2) + 46\,(7) + 41\,(10) + 29\,(30) \\ &= 49 + 96 + 322 + 410 + 870 = 1747 \end{aligned}$$

$$\begin{aligned} x_{Ek} &= \sum_{i=1}^{5}\left[\frac{S_k(i)}{S_k}\right]x_i = \left(\frac{49}{213}\right)1 + \left(\frac{48}{213}\right)2 + \left(\frac{46}{213}\right)7 + \left(\frac{41}{213}\right)10 + \left(\frac{29}{213}\right)30 \\ &= \frac{49 + 96 + 322 + 410 + 870}{213} = \frac{1747}{213} = 8.202 \end{aligned}$$

The average income $\overline{x}$ is equal to $(1 + 2 + 7 + 10 + 30)/5 = 10$ so that the inequality index I_B we obtain is equal to $\left(1 - \frac{8.202}{10}\right) = 0.1798$.

6.1 Breakdown by income sources

Let us go back to the original income distribution $\{1, 2, 7, 10, 30\}$ and assume that there are two income sources. The following table gives the information on the income sources that each individual receives.

	Individual 1	Individual 2	Individual 3	Individual 4	Individual 5
Income source 1	0	1	5	3	20
Income source 2	1	1	2	7	10
Total income	1	2	7	10	30

As shown previously, with $k = 1$, the vector of the values $f_k(i)$ will be: 1, 1, 2, 3, 5 so that the vector of the values $S_k(i)$ will be: 12, 11, 10, 8, 5 while the value of S_k will be 46.

Therefore, the vector of the values $p_k = \left(\frac{f_k(i)}{S_k}\right)$ will be: $\left\{\left(\frac{1}{46}\right)\left(\frac{1}{46}\right), \left(\frac{2}{46}\right), \left(\frac{3}{46}\right), \left(\frac{5}{46}\right)\right\}$ while the row vector $r' = \left\{\left(\frac{x_1}{\bar{x}}\right), \ldots, \left(\frac{x_i}{\bar{x}}\right), \ldots, \left(\frac{x_{nm}}{\bar{x}}\right)\right\} = \left\{\left(\frac{1}{10}\right), \left(\frac{2}{10}\right), \left(\frac{7}{10}\right), \left(\frac{10}{10}\right), \left(\frac{30}{10}\right)\right\}$.

Using (19) we derive that

$$G_k = p_k A[e - r] = \left\{\left(\frac{1}{46}\right)\left(\frac{1}{46}\right), \left(\frac{2}{46}\right), \left(\frac{3}{46}\right), \left(\frac{5}{46}\right)\right\} \begin{bmatrix} 10000 \\ 11000 \\ 11100 \\ 11110 \\ 11111 \end{bmatrix} \begin{Bmatrix} 1 - \left(\frac{1}{10}\right) \\ 1 - \left(\frac{2}{10}\right) \\ 1 - \left(\frac{7}{10}\right) \\ 1 - \left(\frac{10}{10}\right) \\ 1 - \left(\frac{30}{10}\right) \end{Bmatrix} = 0.274$$

For the first income source whose average is (0 + 1 + 5 + 3 + 20)/5 = (29/5) = 5.8, the index will be expressed as

$$G_k^1 = \left\{\left(\frac{1}{46}\right)\left(\frac{1}{46}\right), \left(\frac{2}{46}\right), \left(\frac{3}{46}\right), \left(\frac{5}{46}\right)\right\} \begin{bmatrix} 10000 \\ 11000 \\ 11100 \\ 11110 \\ 11111 \end{bmatrix} \begin{Bmatrix} 1 - \left(\frac{0}{5.8}\right) \\ 1 - \left(\frac{1}{5.8}\right) \\ 1 - \left(\frac{3}{5.8}\right) \\ 1 - \left(\frac{5}{5.8}\right) \\ 1 - \left(\frac{20}{5.8}\right) \end{Bmatrix} = 0.321$$

For the second income source whose average is (1 + 1 + 2 + 7 + 10)/5 = 21/5 = 4.2, the index will be expressed as

$$G_k^2 = \left\{\left(\frac{1}{46}\right)\left(\frac{1}{46}\right), \left(\frac{2}{46}\right), \left(\frac{3}{46}\right), \left(\frac{5}{46}\right)\right\} \begin{bmatrix} 10000 \\ 11000 \\ 11100 \\ 11110 \\ 11111 \end{bmatrix} \begin{Bmatrix} 1 - \left(\frac{1}{4.2}\right) \\ 1 - \left(\frac{1}{4.2}\right) \\ 1 - \left(\frac{2}{4.2}\right) \\ 1 - \left(\frac{7}{4.2}\right) \\ 1 - \left(\frac{10}{4.2}\right) \end{Bmatrix} = 0.229$$

Since the ranking of the second income source is identical to the ranking of total income, for this income source the index PG_k^2 is identical to the index G_k^2.

The ranking of the first income source is however different from the ranking of total income and hence the index PG_k^1 of this source will be different from the index G_k^1. More precisely we can write that

$$PG_k^1 = \left\{\left(\frac{1}{46}\right)\left(\frac{1}{46}\right),\left(\frac{2}{46}\right),\left(\frac{3}{46}\right),\left(\frac{5}{46}\right)\right\}\begin{bmatrix}1\,0\,0\,0\,0\\1\,1\,0\,0\,0\\1\,1\,1\,0\,0\\1\,1\,1\,1\,0\\1\,1\,1\,1\,1\end{bmatrix}\left\{\begin{array}{c}1-\left(\frac{0}{5.8}\right)\\1-\left(\frac{1}{5.8}\right)\\1-\left(\frac{5}{5.8}\right)\\1-\left(\frac{3}{5.8}\right)\\1-\left(\frac{20}{5.8}\right)\end{array}\right\} = 0.307$$

It is then easy to check that $s_1 PG_k^1 + s_2 G_k^2 = \left(\left(\frac{29}{50}\right)0.307\right) + \left(\left(\frac{21}{50}\right)0.229\right) = 0.274$ which is the value of the Gini index of total income, as mentioned previously.

6.2 Decomposition by income classes

Let us go back to the previous illustration where with $k = 1$, the vector of the values $f_k(i)$ is 1, 1, 2, 3, 5 and the vector of the values $p_k = \left(\frac{f_k(i)}{S_k}\right)$ is $\left\{\left(\frac{1}{46}\right)\left(\frac{1}{46}\right),\left(\frac{2}{46}\right),\left(\frac{3}{46}\right),\left(\frac{5}{46}\right)\right\}$.

Let us also define the matrix A as previously. Finally, using the previous illustration, let us define two adjacent income classes. The first one includes the incomes $\{1, 2, 7\}$ and the second one the incomes $\{10, 30\}$. The mean m_1 of the first income class is hence (10/3) while the mean m_2 of the second is 20. The overall mean m is 10.

Using (26) the between groups index will be expressed as

$$p_k A\left[\frac{\overline{x}-\overline{x}_m}{\overline{x}}\right] = \left\{\left(\frac{1}{46}\right)\left(\frac{1}{46}\right),\left(\frac{2}{46}\right),\left(\frac{3}{46}\right),\left(\frac{5}{46}\right)\right\}\begin{bmatrix}1\,0\,0\,0\,0\\1\,1\,0\,0\,0\\1\,1\,1\,0\,0\\1\,1\,1\,1\,0\\1\,1\,1\,1\,1\end{bmatrix}\left\{\begin{array}{c}\left(\frac{10-\left(\frac{10}{3}\right)}{10}\right)\\\left(\frac{10-\left(\frac{10}{3}\right)}{10}\right)\\\left(\frac{10-\left(\frac{10}{3}\right)}{10}\right)\\\left(\frac{10-20}{10}\right)\\\left(\frac{10-20}{10}\right)\end{array}\right\}$$

$$= \left\{\left(\frac{12}{46}\right)\left(\frac{11}{46}\right)\left(\frac{10}{46}\right)\left(\frac{8}{46}\right)\left(\frac{5}{46}\right)\right\}\left\{\begin{array}{c}\left(\frac{20}{30}\right)\\\left(\frac{20}{30}\right)\\\left(\frac{20}{30}\right)\\\left(\frac{-10}{10}\right)\\\left(\frac{-10}{10}\right)\end{array}\right\}$$

$$= \frac{240+220+200-240-150}{1380} = \frac{270}{1380} = 0.196$$

Using again (26) the within groups index will be expressed as

$$p_k A\left[\frac{\overline{x}_m - x_i}{\overline{x}}\right] = \left\{\left(\frac{1}{46}\right)\left(\frac{1}{46}\right), \left(\frac{2}{46}\right), \left(\frac{3}{46}\right), \left(\frac{5}{46}\right)\right\} \begin{bmatrix} 1\,0\,0\,0\,0 \\ 1\,1\,0\,0\,0 \\ 1\,1\,1\,0\,0 \\ 1\,1\,1\,1\,0 \\ 1\,1\,1\,1\,1 \end{bmatrix} \left\{\begin{matrix} \left(\frac{\left(\frac{10}{3}\right)-1}{10}\right) \\ \left(\frac{\left(\frac{10}{3}\right)-2}{10}\right) \\ \left(\frac{\left(\frac{10}{3}\right)-7}{10}\right) \\ \left(\frac{20-10}{10}\right) \\ \left(\frac{20-30}{10}\right) \end{matrix}\right\}$$

$$= \left\{\left(\frac{12}{46}\right)\left(\frac{11}{46}\right)\left(\frac{10}{46}\right)\left(\frac{8}{46}\right)\left(\frac{5}{46}\right)\right\} \left\{\begin{matrix} \left(\frac{7}{30}\right) \\ \left(\frac{4}{30}\right) \\ \left(\frac{-11}{30}\right) \\ \left(\frac{30}{30}\right) \\ \left(\frac{-30}{30}\right) \end{matrix}\right\}$$

$$= \frac{84 + 44 - 110 + 240 - 150}{1380} = \frac{368 - 260}{1380} = \frac{108}{1380} = 0.078$$

The sum of the between and within groups indices is then equal to $(0.196+0.078) = 0.274$ i.e.d.

6.3 Decomposition by population subgroups

Let us assume that there are two population subgroups in the population, group *A* and group *B*. Here is the affiliation of the various individuals to these two groups.

	Individual 1	Individual 2	Individual 3	Individual 4	Individual 5
Total income	1	2	7	10	30
Group to which the individual belongs	*A*	*B*	*A*	*B*	*A*

The average income in group A is $(1 + 7 + 30)/3 = (38/3) = 12.666$ and the average income in group B is $(2 + 10)/2 = 6$.

The vector x of the incomes ranked by non-decreasing values will then be expressed as

$$x = \{1, 2, 7, 10, 30\}$$

while the vector y of these same incomes ranked first by non-decreasing values of the average incomes of the population subgroups to which they belong, and then within each population subgroups by non-decreasing values of the incomes of the individuals who belong to the corresponding population subgroup, will be expressed as

$$y = \{2, 10, 1, 7, 30\}.$$

As shown in (29), the between groups index G_B is written as

$$G_B = p_k A\left[\frac{(\overline{y} - \overline{y}_m)}{\overline{y}}\right]$$

$$G_B = \left\{\left(\frac{1}{46}\right)\left(\frac{1}{46}\right), \left(\frac{2}{46}\right), \left(\frac{3}{46}\right), \left(\frac{5}{46}\right)\right\}\begin{bmatrix} 1\,0\,0\,0\,0 \\ 1\,1\,0\,0\,0 \\ 1\,1\,1\,0\,0 \\ 1\,1\,1\,1\,0 \\ 1\,1\,1\,1\,1 \end{bmatrix}\left\{\begin{matrix} \left(\frac{10-6}{10}\right) \\ \left(\frac{10-6}{10}\right) \\ \left(\frac{10-12.666}{10}\right) \\ \left(\frac{10-12.666}{10}\right) \\ \left(\frac{10-12.666}{10}\right) \end{matrix}\right\}$$

$$G_B = \left\{\left(\frac{12}{46}\right)\left(\frac{11}{46}\right)\left(\frac{10}{46}\right)\left(\frac{8}{46}\right)\left(\frac{5}{46}\right)\right\}\left\{\begin{matrix} \left(\frac{4}{10}\right) \\ \left(\frac{4}{10}\right) \\ \left(\frac{-2.666}{10}\right) \\ \left(\frac{-2.666}{10}\right) \\ \left(\frac{-2.666}{10}\right) \end{matrix}\right\}$$

$$G_B = \frac{48 + 44 - 26.66 - 21.328 - 13.33}{460} = \frac{30.682}{460} = 0.0667$$

Similarly, as shown in (29), the within groups index G_W is expressed as

$$G_W = p_k A\left[\frac{(\overline{y}_m - y_i)}{\overline{y}}\right]$$

$$G_W = \left\{\left(\frac{1}{46}\right)\left(\frac{1}{46}\right), \left(\frac{2}{46}\right), \left(\frac{3}{46}\right), \left(\frac{5}{46}\right)\right\}\begin{bmatrix} 1\,0\,0\,0\,0 \\ 1\,1\,0\,0\,0 \\ 1\,1\,1\,0\,0 \\ 1\,1\,1\,1\,0 \\ 1\,1\,1\,1\,1 \end{bmatrix}\left\{\begin{matrix} \left(\frac{6-2}{10}\right) \\ \left(\frac{6-10}{10}\right) \\ \left(\frac{12.666-1}{10}\right) \\ \left(\frac{12.666-7}{10}\right) \\ \left(\frac{12.666-30}{10}\right) \end{matrix}\right\}$$

$$G_W = \left\{\left(\frac{12}{46}\right)\left(\frac{11}{46}\right)\left(\frac{10}{46}\right)\left(\frac{8}{46}\right)\left(\frac{5}{46}\right)\right\}\left\{\begin{matrix} \left(\frac{4}{10}\right) \\ \left(\frac{-4}{10}\right) \\ \left(\frac{11.666}{10}\right) \\ \left(\frac{5.666}{10}\right) \\ \left(\frac{-17.334}{10}\right) \end{matrix}\right\}$$

$$G_W = \frac{48 - 44 + 116.66 + 45.328 - 86.67}{460} = \frac{79.318}{460} = 0.172$$

Finally, the overlapping term OV will be written as

$$OV = \left\{\left(\frac{1}{46}\right)\left(\frac{1}{46}\right), \left(\frac{2}{46}\right), \left(\frac{3}{46}\right), \left(\frac{5}{46}\right)\right\}\begin{bmatrix} 1\,0\,0\,0\,0 \\ 1\,1\,0\,0\,0 \\ 1\,1\,1\,0\,0 \\ 1\,1\,1\,1\,0 \\ 1\,1\,1\,1\,1 \end{bmatrix}\left[\left\{\begin{matrix} \left(\frac{2}{10}\right) \\ \left(\frac{10}{10}\right) \\ \left(\frac{1}{10}\right) \\ \left(\frac{7}{10}\right) \\ \left(\frac{30}{10}\right) \end{matrix}\right\} - \left\{\begin{matrix} \left(\frac{1}{10}\right) \\ \left(\frac{2}{10}\right) \\ \left(\frac{7}{10}\right) \\ \left(\frac{10}{10}\right) \\ \left(\frac{30}{10}\right) \end{matrix}\right\}\right]$$

Table 1 Single-parameter generalization of the Gini index [10]

	$k = 0$	$k = 1$	$k = 2$	$k = 3$
G_k	0.247	0.050	0.025	0.016

$$OV = \left\{\left(\frac{12}{46}\right)\left(\frac{11}{46}\right)\left(\frac{10}{46}\right)\left(\frac{8}{46}\right)\left(\frac{5}{46}\right)\right\}\left\{\begin{matrix}\left(\frac{1}{10}\right)\\ \left(\frac{8}{10}\right)\\ \left(\frac{-6}{10}\right)\\ \left(\frac{-3}{10}\right)\\ \left(\frac{0}{10}\right)\end{matrix}\right\}$$

$$OV = \frac{12 + 88 - 60 - 24 + 0}{460} = \frac{16}{460} = 0.035$$

The sum of the between index, the within index and the overlapping term is then expressed as $(0.0667 + 0.172 + 0.035) = 0.2737 \cong 0274$ i.e.d.

7 Inequality in expenditure on social protection among European countries

In this section, we illustrate the applicability of the previous formulations for the analysis of inequality among the European Union (EU-28) countries, as far as expenditures on social protection in the year 2018 are concerned. The data utilized for this analysis is sourced from Eurostat, specifically derived from the European System of Integrated Social Protection Statistics (ESSPROS). ESSPROS constitutes a standardized framework that facilitates the cross-national evaluation of domestically collected data pertaining to social protection. The data extraction process was executed as of the 18th of August 2023, with reference to the year 2018. This temporal choice was made to mitigate the influence of the pandemic on social protection expenditures. Note also that the data is standardized and expressed in terms of Parity Purchasing Standards[2] (PPS) per capita.

Average expenditure on social protection in EU-28 in 2018 is 6712.13 PPS per capita but this expenditure is not homogeneous across countries, nor among welfare systems, not within member states systems, and not even among functions of social protections. In what follows we analyze inequality in social protection expenditure, disaggregated by functions and by welfare systems, providing results for different income weights derived from Fibonacci-like sequences of numbers.

First, we analyze inequality on social protection expenditure among EU-28 countries for different values of the parameter k. The greater the value of the parameter, the higher the weight given to lower incomes, and when $k = 0$, the inequality index is related to the Gini index (Table 1).

Inequality decreases as the value of the parameter increases, meaning that as we give more importance to countries with lower expenditure on social protection inequality is lower. At the same time G_0 differs from the Gini index (0.256), as explained in the appendix, where: $G = G_0\left(1 + \frac{1}{28}\right) = 0.247\left(1 + \frac{1}{28}\right) = 0.256$.

---

[2] Purchasing power standard, is an artificial currency unit. Theoretically, one PPS can buy the same amount of goods and services in each country. However, price differences across borders mean that different amounts of national currency units are needed for the same goods and services depending on the country.

Table 2 Single-parameter generalization of the Gini index and contribution by function

		$k = 0$	$k = 1$	$k = 2$	$k = 3$
G_k	Overall	0.247	0.050	0.025	0.016
	Sickness/health care	0.256	0.053	0.025	0.016
	Disability	0.361	0.098	0.050	0.032
	Old age	0.225	0.039	0.018	0.011
	Survivors	0.393	0.088	0.045	0.030
	Family/children	0.318	0.093	0.050	0.034
	Unemployment	0.385	0.073	0.033	0.020
	Housing	0.589	0.139	0.066	0.042
	Social exclusion	0.479	0.118	0.059	0.038
PG_k	Sickness/health care	0.237	0.044	0.021	0.013
	Disability	0.331	0.095	0.050	0.032
	Old age	0.216	0.036	0.017	0.011
	Survivors	0.225	0.056	0.033	0.024
	Family/children	0.274	0.090	0.050	0.034
	Unemployment	0.316	0.043	0.018	0.011
	Housing	0.375	0.036	0.009	0.003
	Social exclusion	0.392	0.077	0.037	0.023
Contribution (%)	Sickness/health care	27.2	24.8	23.6	23.3
	Disability	11.0	15.5	16.4	16.4
	Old age	36.1	29.5	27.8	27.2
	Survivors	4.6	5.6	6.7	7.3
	Family/children	10.3	16.6	18.6	19.4
	Unemployment	5.0	3.4	2.9	2.7
	Housing	2.0	0.9	0.5	0.2
	Social exclusion	3.8	3.7	3.5	3.4

Social benefits can be classified by function. The latter indicates the primary purpose for which social protection is provided. Eight functions of social protection are distinguished in the ESSPROS:

1. Sickness/health care
2. Disability
3. Old age
4. Survivors
5. Family/children
6. Unemployment
7. Housing
8. Social exclusion not elsewhere classified.

We analyze inequality in each of the functions and the decomposition of overall inequality in social expenditure by functions.

Table 2 presents the contribution of each function m to the overall inequality: $Contribution\ of\ m = \frac{\left(\frac{\overline{x}_m}{\overline{x}}\right) G_{k,m} GC_{k,m}}{\sum_{m=1}^{M}\left(\frac{\overline{x}_m}{\overline{x}}\right) G_{k,m} GC_{k,m}}$. Notably, regardless of the parameter used, the

Table 3 Single-parameter generalization of the Gini index by Welfare system and decomposition by subgroups

		$k = 0$	$k = 1$	$k = 2$	$k = 3$
G_k	Overall	0.2475	0.0502	0.0248	0.0161
	Nordic	0.026	0.017	0.012	0.009
	Conservative	0.020	0.020	0.012	0.009
	Liberals	0.064	0.039	0.025	0.018
	Southern	0.092	0.053	0.033	0.023
	Central and Eastern Europe	0.105	0.041	0.022	0.014
G_{Bk}		0.2233	0.0380	0.0164	0.0100
G_{Wk}		0.0202	0.0109	0.0080	0.0060
$\text{Overlapping}_\text{k}$		0.0040	0.0013	0.0004	0.0002

housing function consistently exhibits the smallest contribution. This is not due to its low level of inequality (in fact, this function has the highest level of inequality) but to the low weight of expenditure on housing in total social expenditure.[3]

Conversely, the old-age function displays a contrasting pattern. Despite having the lowest level of inequality, its contribution to overall inequality is the most substantial because of its high weight in total social expenditure (see Table 4 in Appendix B). Following closely is the sickness/health care function.

Furthermore, note that when we increase the weight given to the lowest expenditures within each function, the contribution of family/children function increases the most. This underscores the fact that inequality is more pronounced at the lower levels of expenditure within this function. Conversely, a different trend emerges for the old-age function, which exhibits higher inequality among countries with the highest levels of expenditure.

In summary, Table 2 provides insights into the varying contributions of functions to the overall inequality. The intricate interplay between inequality and expenditure across different functions underscores the nuanced nature of these relationships.

We anticipate that welfare systems will exhibit variations based on levels of social protection expenditure. To address this, our assessment focuses on both inter-system and intra-system inequalities. We group countries according to the classification of welfare typologies introduced by Ferrera [5] and Mensah and Adjei [7]: Nordic (Sweden, Denmark, Finland, and Norway), Conservative (Austria, Belgium, France, Germany, Netherland, Luxembourg, and Switzerland), Liberals (United Kingdom and Ireland), Southern Europe (Greece, Spain, Italy, Portugal, Cyprus, and Malta), and Central and Eastern Europe (Estonia, Lithuania, Hungary, Czech Republic, Poland, Latvia, Romania, Slovakia, Slovenia, Bulgaria, and Croatia).

Table 3 reports the inequality levels within each welfare system and the within, between and overlapping components of inequality.

As anticipated, the inequality observed between different welfare systems is more pronounced than the inequality within welfare systems. The classification of welfare systems is based on various attributes, the level of expenditure being an important attribute. This pattern is observed regardless of the parameter utilized. While there is a certain degree of overlap among these welfare systems, the extent of such overlap remains limited.

---

[3] See Appendix B for a table that illustrates the proportion that the average expenditure per function represents in relation to the total expenditure on social benefits.

Furthermore, as we allocate greater weight to the lower levels of expenditure, the relative contribution of inequality between welfare systems decreases. In contrast, the contribution of inequality within welfare systems becomes more prominent. This highlights the fact that the lowest expenditure levels across various welfare systems exhibit a higher degree of uniformity. Conversely, within each specific welfare system, the lowest expenditure levels are characterized by greater inequality compared to the overall scenario.

This comprehensive analysis offers valuable insights into the disparities present in social protection expenditure among different countries. These insights hold the potential to inform the development of cohesive policies within the European Union.

8 Conclusions

Following recent work of Subramanian [10] and Subramanian and Creedy [2] on a generalization of the Gini index based on the 'metallic' sequences of number theory, this paper showed how this new generalization of the Gini index could be decomposed by income sources, income classes and population subgroups. While the traditional Gini index may be also decomposed by income sources and classes and by population subgroups, the generalization of the Gini index based on the 'metallic' sequences of number theory, has the advantage of offering a breakdown of an inequality index which gives a greater weight to lower incomes. An empirical illustration then applied this new extension of the Gini index to an analysis of social protection expenditure per capita in the EU-28 in 2018. It turns out that this expenditure varied significantly across countries, welfare systems, member states, and social protection functions. We assessed inequality across functions and welfare systems using different income weights derived from Fibonacci-like sequences.

It appears that the housing function consistently exhibits the smallest contribution, due to its low expenditure weight, while the old-age function has the lowest inequality but the highest contribution, due to its substantial expenditure weight. The family/children function's contribution increases when a higher weight is given to lower expenditures, highlighting a pronounced inequality at lower expenditure levels. Conversely, the old-age function shows higher inequality among countries with higher expenditures.

We also explored variations in welfare systems based on social protection expenditure levels, observing greater inequality between systems than within them. These findings offer valuable insights for informing cohesive policies within the European Union.

Appendices

Appendix A: The link between the index G_0 and the traditional Gini index

Creedy and Subramanian [1] defined a Borda type of welfare function W_B as

$$W_B = \sum_{i=1}^{n} (n+1-i)x_i \tag{30}$$

whose "equally distributed equivalent level of income" x_{EB} will be expressed as

$$x_{EB} = \sum_{i=1}^{n} \frac{(n+1-i)}{\sum_{i=1}^{n}(n+1-i)} x_i \tag{31}$$

The corresponding inequality index I_B will then be written as

$$I_B = 1 - \frac{\left[\sum\limits_{i=1}^{n} \frac{(n+1-i)}{\sum\limits_{i=1}^{n}(n+1-i)} x_i\right]}{\overline{x}} \tag{32}$$

Since $\sum\limits_{i=1}^{n}(n+1-i) = n\left(\frac{n+1}{2}\right)$, we derive that

$$I_B = 1 - \left\{\left[\sum_{i=1}^{n}(n+1-i)\left(\frac{x_i}{\overline{x}}\right)\right]\left[\frac{2}{n(n+1)}\right]\right\} \tag{33}$$

The traditional Gini index I_G may be expressed (see, [1] as

$$I_G = \left(\frac{n+1}{n}\right) - \left\{\left[\left(\frac{2}{n^2}\right)\right]\left[\sum_{i=1}^{n}(n+1-i)\left(\frac{x_i}{\overline{x}}\right)\right]\right\} \tag{34}$$

and

$$I_G = 1 + \left(\frac{1}{n}\right) - \left\{\left(\frac{2}{n^2}\right)\left[\sum_{i=1}^{n}(n+1-i)\left(\frac{x_i}{\overline{x}}\right)\right]\right\} \tag{35}$$

Combining (33) and (35) we then write that

$$\left[\sum_{i=1}^{n}(n+1-i)\left(\frac{x_i}{\overline{x}}\right)\right] = (1 - I_B)\left[n\left(\frac{n+1}{2}\right)\right] = \left(1 + \frac{1}{n} - I_G\right)\left(\frac{n^2}{2}\right) \tag{36}$$

so that

$$\begin{aligned} \frac{n^2}{2} + \frac{n}{2} - I_B\left(\frac{n^2}{2} + \frac{n}{2}\right) &= \frac{n^2}{2} + \frac{n}{2} - I_G\left(\frac{n^2}{2}\right) \\ I_B\left(\frac{n^2}{2} + \frac{n}{2}\right) &= I_G\left(\frac{n^2}{2}\right) \\ I_G &= I_B\left(1 + \frac{1}{n}\right) \end{aligned} \tag{37}$$

For small values of n there will be a difference between I_G and I_B, but for values of n sufficiently big the values of the indices I_G and I_B will be more or less the same.

Note finally that index I_B is in fact identical to the index G_0 defined in (10) when $k = 0$, and that in such a case the elements of the vector $S_k(i)$ defined in (2) are in fact the Borda numbers.

Appendix B

See Table 4.

Table 4 Total expenditure in social benefits and proportion that the average expenditure per function represents in relation to the total expenditure on social benefits

Total expenditure on social benefits	6712.133
Sickness/health care	28%
Disability	8%
Old age	41%
Survivors	5%
Family/children	9%
Unemployment	4%
Housing	1%
Social exclusion	2%

Funding Funding for open access publishing: Universidad Málaga/CBUA.

Data availability The data sources are explicitly mentioned in the first paragraph of Section 7. More precisely the data on ESSPROS (European system of integrated social protection statistics) are publicly available on the web at https://ec.europa.eu/eurostat/web/social-protection/database.

References

1. Creedy, J., Subramanian, S.: Exploring a new class of inequality measures and associated value judgements: Gini and Fibonacci-type sequences. Working Paper 11/2022, Working Papers in Public Finance, Victoria Business School, Victoria University of Wellington, New Zealand (2022)
2. Creedy, J., and S. Subramanian.: Exploring a new class of inequality measures and associated value judgements: Gini and Fibonacci-type Sequences. Sankhyā: The Indian Journal of Statistics **85B**(1), 110-131 (2023)
3. Donaldson, D., Weymark, J.A.: A single-parameter generalization of the Gini indices of inequality. J. Econ. Theory **22**(1), 67–86 (1980)
4. Fei, J.C.H., Ranis, G., Kuo, S.W.Y.: Growth and the family distribution of income by factor components. Quart. J. Econ. **92**, 17–53 (1978)
5. Ferrera, M.: The 'Southern model' of welfare in social Europe. J. Eur. Soc. Policy **6**, 17–37 (1996)
6. Giorgi, G.M.: Bibliographic portrait of the Gini concentration ratio. Metron XLVIII **1–4**, 183–221 (1990)
7. Mensah, A., Adjei, N.K.: Work-life balance and self-reported health among working adults in Europe: a gender and welfare state regime comparative analysis. BMC Public Health **20**, 1052 (2020)
8. Rao, V.N.: Two decompositions of concentration ratios. J. R. Stat. Soc. Ser. A **132**, 418–425 (1969)
9. Silber, J.: Factor components, population subgroups and the computation of the Gini index. Rev. Econ. Stat. **71**(1), 107–115 (1989)
10. Subramanian, S.: A single-parameter generalization of Gini based on the 'metallic' sequences of number theory. Econ. Bull. **41**(4), 2309–2319 (2021)
11. Xu, K.: How has the literature on Gini's index evolved in the past 80 Yars?. Department of Economics, Dalhousie University, Canada. The Chinese version of this paper was published in the *China Economic Quarterly* (2003), vol. 2, pp. 757–778 (2004)
12. Yitzhaki, S.: On an extension of the Gini inequality index. Int. Econ. Rev. **24**(3), 617–628 (1983)