Measuring income inequality based on unequally distributed income

Joongyang Park¹ · Youngsoon Kim¹ · Ae-Jin Ju²

Received: 21 August 2019 / Accepted: 4 August 2020 / Published online: 13 August 2020
© The Author(s) 2020

Abstract

This paper proposes a new framework for measuring income inequality. The framework is based on the unequally distributed (UD) incomes that are obtained by removing the equally distributed parts from incomes. We then derive the normalized norm indexes from the cumulative distribution function and the unscaled Lorenz curve of the UD incomes. The relation between the normalized norm indexes and the popular Gini coefficient and coefficient of variation (CV) shows that the Gini coefficient and CV represent only parts of income inequality. We analyze example income distributions and the Luxembourg Income Study datasets to show that the normalized norm indexes evaluate income inequality appropriately and solve the negative income problem.

Keywords Coefficient of variation · Gini coefficient · Income inequality · ℓ_1 Norm · ℓ_2 Norm · Normalized norm index · Rawlsian index · Unequally distributed income

JEL Classification C43 · D31 · D63

1 Introduction

Economic inequality is a major topic in economic research. Inequalities in economic resources and factors are useful to understand economic processes. Income inequality is central to issues related to economic inequality. Since Lorenz (1905) introduced the Lorenz curve, various income inequality indexes have been developed to quantitatively measure the degree of income inequality. Such indexes are used to examine

✉ Joongyang Park
joongyang.park@gmail.com

¹ Department of Information and Statistics, Gyeongsang National University, 501 Jinju-daero, Jinju-si, Gyeongsangnam-do 52828, Republic of Korea

² Department of Information and Statistics, Graduate School, Gyeongsang National University, Jinju 52828, Republic of Korea

whether income inequality is decreasing, whether a redistribution policy is effective at reducing economic inequality, whether a country is more or less unequal than other countries, and so on. The conventional approach is to compute a single index from the income distribution and the corresponding Lorenz curve. The Gini coefficient, defined statistically as the mean absolute difference of all pairs of incomes divided by twice the mean income, was introduced by Gini (1914). The Gini coefficient was further generalized by Donaldson and Weymark (1980) and Yitzhaki (1983). Gastwirth (2014) showed that the Gini coefficient underestimates the rate increase in inequality and modified the Gini coefficient by dividing the mean absolute difference of all pairs of incomes by the median income. Other measures include the coefficient of variation (CV) and the Pietra index. The Pietra index suggested by Schutz (1951) is the portion of the total income that would have to be redistributed. The Pietra index is also known as the Hoover index, the Robin Hood index, and the Ricci-Schutz index. Zanardi (1964) developed an index to measure the asymmetry of the Lorenz curve. In recent contributions, Gallegati et al. (2016) and Clementi et al. (2019) propose the Zanardi index as an appropriate measure of income inequality. There is a class of indexes based on information theory, members of which include the Atkinson index, the generalized entropy index, and the Theil index (Atkinson 1970; Shorrocks 1980; Theil 1967). Such indexes measure diversity in incomes. For a general overview of income inequality measurement, refer to Hao and Naiman (2010), Jenkins and Kerm (2011) and Cowell (2011).

As mentioned above, most of the currently available income inequality indexes are derived from the income distribution and the corresponding Lorenz curve. This approach suffers from the negative income problem. Negative incomes are frequently observed in reality. Self-employed individuals will have negative incomes when their firms and organizations experience losses. Negative incomes are also observed when employees repay an amount of debt that exceeds their earnings. Since most income inequality indexes assume that incomes are nonnegative, the negative incomes observed in the reality create problems in inequality measurement. The Atkinson index, the generalized entropy index, and the Theil index are not computable. The Gini coefficient requires the normalization proposed by Chen et al. (1982) and Raffinetti et al. (2015). It is well known that the CV and Pietra index are appropriate only for nonnegative variables at a ratio scale. Since the income shares for negative incomes cannot be interpreted as proportions, all income inequality indexes involving income shares are neither computable nor interpretable. Therefore, various remedial adjustments of negative incomes have been devised. For example, Cowell (2011) and Hao and Naiman (2010) suggest the deletion of non-positive incomes. OECD (2016) recommends replacing negative incomes with zero incomes. Bellù and Liberati (2006) advise replacing zero incomes with arbitrarily small positive incomes. Raffinetti et al. (2016) study the effects of these remedial adjustments on the Gini coefficient.

These remedial adjustments indicate that the nonnegative income assumption does not correspond to reality. The conventional approach attempts to resolve the inconsistency between this assumption and reality by adjusting the data, not the assumption. Therefore, we need a realistic approach based on something other than the income distribution.

In a recent work, Park et al. (2018) suggest that income inequality should be measured from unequally distributed (UD) incomes, not incomes. UD incomes are obtained from the income distribution by removing the parts of incomes associated with equality. The UD income approach neither requires the nonnegative income assumption nor suffers from the negative income problem. Park et al. (2018) also propose that income inequality should be represented by two indexes reflecting the location and dispersion of UD incomes. However, they do not present a general framework for measuring income inequality based on UD incomes. The aims of the present paper are (i) to propose a framework using the cumulative distribution function (CDF) and the unscaled Lorenz curve of UD incomes; (ii) to show that income inequality has two dimensions, horizontal and vertical, under the proposed framework; (iii) to develop new indexes for the two dimensions; and (iv) to present the relation between the new indexes and the popular Gini coefficient and CV.

This paper is organized as follows. Section 2 briefly reviews UD income. Section 3 adopts the CDF of UD income and represents income inequality as two types of departure from perfect equality. Normalized norms of the departures are proposed as new income inequality indexes. It is also shown that the normalized norm indexes are decomposed into three constituents, the Rawlsian index, the Gini coefficient, and the CV. Section 4 adopts the unscaled Lorenz curve of UD income and develops an income inequality index. It is shown that the index is equivalent to one of the normalized norm indexes derived in Sect. 3. The normalized norm indexes are illustrated using example income distributions and the Luxembourg Income Study (LIS) datasets in Sect. 5. Section 6 presents concluding remarks.

2 Unequally distributed income

We denote the income distribution of a population of n individuals by $\mathbf{y} = (y_1, y_2, \dots, y_n)$, where y_i is the income of i -th individual. Without loss of generality, we assume that the incomes are ordered, that is, $y_1 \leq y_2 \leq \dots \leq y_n$. The total income and mean income are denoted by $S_y = \sum_{i=1}^n y_i$ and $\mu_y = S_y/n$, respectively.

We begin by noting that $y_i = y_1 + (y_i - y_1)$, $i = 1, 2, \dots, n$. This expression indicates that all income values include y_1 , ny_1 of S_y is equally distributed among the n individuals, and the remaining $(S_y - ny_1)$ is unequally distributed among the n individuals. The unequally distributed portions of the incomes, $x_i = (y_i - y_1)$, $i = 1, 2, \dots, n$, are referred to as the UD incomes. The total and mean of the UD incomes are written as

$$S_x = (S_y - ny_1) = n(\mu_y - y_1) \text{ and } \mu_x = (\mu_y - y_1).$$

We denote the UD income distribution by $\mathbf{x} = (x_1, x_2, \dots, x_n)$.

The income distribution encompasses the UD income distribution, and contains information about both equality and inequality. The UD income distribution is obtained by removing the parts of the incomes associated with equality from the income distribution. The amount of income associated with equality is ny_1 . The amount of income

associated with inequality is $(S_y - ny_1)$, which is distributed as $\mathbf{x}$. Income inequality definitely exists in the UD income distribution. It is therefore reasonable to measure income inequality from the UD income distribution.

As defined in Cowell (2011), inequality is a departure from some idea of equality. In the field of income inequality, equality is represented by a perfectly equal income distribution $\mathbf{y}_{pe} = (\mu_y, \dots, \mu_y)$, usually referred to as perfect equality. We should note that the most significant change from using the UD income approach occurs in the representation of perfect equality. The UD income distribution corresponding to perfect equality is obtained as $\mathbf{x}_{pe} = (0, 0, \dots, 0)$, the origin of the UD income space.

The income inequality of an income distribution is usually evaluated by the discrepancy from its corresponding $\mathbf{y}_{pe}$. Two income distributions with the same n and different S_y correspond to different $\mathbf{y}_{pe}$ s. Theoretically, there are infinitely many $\mathbf{y}_{pe}$ s in the income space. Each of the two income distributions is compared to its corresponding $\mathbf{y}_{pe}$. Two income distributions with the same n and S_y in general have different ny_1 s. In this case, the two income distributions are compared to the same $\mathbf{y}_{pe}$. However, regardless of S_y and μ_y , perfect equality in the UD income approach is uniquely represented by $\mathbf{x}_{pe}$, the origin of the UD income space. The following two sections address how to measure the discrepancy of $\mathbf{x}$ from $\mathbf{x}_{pe}$.

3 Normalized norm indexes from CDF

A general way to describe a distribution is the CDF. The CDFs for $\mathbf{x}$ and $\mathbf{x}_{pe}$ are

$$F(x) = \begin{cases} 0 & \text{for } x < x_1, \\ \frac{i}{n} & \text{for } x_i \leq x < x_{i+1}, i = 1, 2, \dots, n-1, \\ 1 & \text{for } x \geq x_n, \end{cases}$$

and

$$F_{pe}(x) = \begin{cases} 0 & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases},$$

respectively. Figure 1 depicts $F(x)$ and $F_{pe}(x)$. Income inequality is then represented by the departure of $F(x)$ from $F_{pe}(x)$. There are two types of departure, the vertical departure and the horizontal departure. The vertical departure is the difference between $F(x)$ and $F_{pe}(x)$. Let $Q(p)$ denote the quantile function of the UD income distribution $\mathbf{x}$, which is defined as

$$Q(p) = \inf \{x : F(x) \geq p\} = \begin{cases} 0 & \text{for } p = 0 \\ x_i & \text{for } \frac{i-1}{n} < p \leq \frac{i}{n}, i = 1, 2, \dots, n. \end{cases}$$

The quantile function for $\mathbf{x}_{pe}$ is obtained as $Q_{pe}(p) = 0$ for $0 \leq p \leq 1$. The horizontal departure is the difference between $Q(p)$ and $Q_{pe}(p)$. The size of the difference between two functions is usually assessed by norms. Thus, we compute the ℓ_1 and ℓ_2 norms of the vertical and horizontal departures.

Fig. 1 Two types of departures of $F(x)$ from $F_{pe}(x)$

Let us first consider the vertical departure. The ℓ_1 and ℓ_2 norms of the vertical departure, denoted by ℓ_1^v and ℓ_2^v , are defined as

$$\ell_1^v = \int (1 - F(x)) dx \text{ and } \ell_2^v = \left[\int (1 - F(x))^2 dx \right]^{\frac{1}{2}}.$$

Applying the results of Dorfman (1979) and Yitzhaki (1998), we have

$$\ell_1^v = \mu_x \text{ and } \ell_2^v = \left(\mu_x - \frac{\Gamma_x}{2} \right)^{\frac{1}{2}}, \quad (1)$$

where Γ_x is the Gini mean difference (GMD) of the UD income. To convert the norms to indexes, we normalize ℓ_1^v and ℓ_2^v . Since ℓ_1^v is in monetary units and ℓ_2^v is in the square root of monetary units, ℓ_1^v is normalized by μ_y , while ℓ_2^v is normalized by $\sqrt{\mu_y}$. The normalized vertical norm indexes, denoted by $\tilde{\ell}_1^v$ and $\tilde{\ell}_2^v$, are obtained as

$$\tilde{\ell}_1^v = \frac{\mu_x}{\mu_y} \text{ and } \tilde{\ell}_2^v = \left(\frac{\mu_x}{\mu_y} - \frac{\Gamma_x}{2\mu_y} \right)^{\frac{1}{2}}. \quad (2)$$

Since $\Gamma_x = \Gamma_y$ and the Gini coefficient is the GMD divided by twice the mean, we have

$$\tilde{\ell}_2^v = \left[\frac{\mu_x}{\mu_y} (1 - G_x) \right]^{\frac{1}{2}} = \left(\frac{\mu_x}{\mu_y} - G_y \right)^{\frac{1}{2}}, \quad (3)$$

where G_y and G_x are the Gini coefficients of income and UD income.

The normalized average vertical distance from perfect equality, $\tilde{\ell}_1^v$, is equal to the Rawlsian index introduced by Park et al. (2018). Since the Rawlsian index is the ratio of total UD income to total income, it can be interpreted as the magnitude of

income inequality. There is another justification for the Rawlsian index. Since the y_i s are ordered, the simplest expression for perfect equality is $y_1 = \mu_y$. Therefore, the discrepancy $\mu_x = (\mu_y - y_1)$ and the normalized μ_x are sensible measures for the departure from perfect equality.

$\tilde{\ell}_2^v$ is obtained as a combination of the Rawlsian index and the Gini coefficient. It should be noted that the Gini coefficient is associated with the vertical departure and does not alone describe income inequality. Equation (3) contradicts the conventional notion that the greater income inequality is, the greater the Gini coefficient is. If a progressive transfer is not directed to the poorest, the mean UD income and the Rawlsian index are preserved while the Gini coefficient is reduced. Equation (3) states that progressive transfers preserving the mean UD income do not decrease the vertical departure. Income inequality can be improved through the progressive transfers reducing the mean UD income.

Next we consider the horizontal departure. As shown in Fig. 1, the horizontal departure is x_i for $(i-1)/n < F(x) \leq i/n$. The ℓ_1 and ℓ_2 norms of the horizontal departure, denoted by ℓ_1^h and ℓ_2^h , are obtained as

$$\ell_1^h = \frac{\sum_{i=1}^n x_i}{n} = \mu_x \text{ and } \ell_2^h = \left(\frac{\sum_{i=1}^n x_i^2}{n} \right)^{\frac{1}{2}} = (\mu_x^2 + V_x)^{\frac{1}{2}}, \quad (4)$$

where V_x is the variance of UD income. Since both ℓ_1^h and ℓ_2^h are in monetary units, we normalize ℓ_1^h and ℓ_2^h by μ_y . The normalized ℓ_1^h and ℓ_2^h are obtained as

$$\tilde{\ell}_1^h = \tilde{\ell}_1^v \text{ and } \tilde{\ell}_2^h = \frac{\mu_x}{\mu_y} (1 + CV_x^2)^{\frac{1}{2}} = \left[\left(\frac{\mu_x}{\mu_y} \right)^2 + CV_y^2 \right]^{\frac{1}{2}}, \quad (5)$$

where CV_y and CV_x are the CVs of income and UD income. It is easily verified that $\tilde{\ell}_2^h$ is equal to the normalized Euclidean distance of $\mathbf{x}$ from perfect equality in the UD income space. Henceforth, we denote $\tilde{\ell}_1^v$ and $\tilde{\ell}_1^h$ simply by $\tilde{\ell}_1$ and call it the Rawlsian index.

We should note that the CV is associated with the horizontal departure. The Gini coefficient and CV are associated with different types of departure. Therefore, they are not substitutes for one another. Progressive transfers preserving the mean UD income reduce the horizontal departure but increase the vertical departure. Such progressive transfers affect the vertical and horizontal departures in opposite directions.

Most conventional income inequality indexes such as G_y and CV_y describe the dispersion of the income distribution. The Rawlsian index measures the location of the UD income distribution, while G_x and CV_x measure the dispersion of the UD income distribution. $\tilde{\ell}_2^v$ and $\tilde{\ell}_2^h$ are expressed as the combinations of the Rawlsian index and a dispersion measure.

The normalized norm indexes $\tilde{\ell}_2^v$ and $\tilde{\ell}_2^h$ suggest the following:

- (i) Income inequality has multiple dimensions. The normalized norm indexes $\tilde{\ell}_2^v$ and $\tilde{\ell}_2^h$ consider two dimensions of income inequality.

- (ii) The normalized norm indexes $\tilde{\ell}_2^v$ and $\tilde{\ell}_2^h$ involve both the location and dispersion of the UD income distribution. The location of the UD income distribution, described by the Rawlsian index $\tilde{\ell}_1$, is the primary constituent of income inequality.
- (iii) The dispersion of the income distribution should be distinguished from income inequality. The popular Gini coefficient and CV are associated with this dispersion. According to Equations (3) and (5), the Gini coefficient and CV represent only parts of income inequality.
- (iv) Progressive transfers preserving the mean UD income reduce dispersion of the UD income distribution. Such a reduction in dispersion decreases the horizontal departure but increases the vertical departure. An effective way to reduce both types of departures is to decrease the Rawlsian index. The Rawlsian index can be decreased by progressive transfers to the poorest.

4 A normalized norm index from the unscaled Lorenz curve

One of the most popular graphical representations of an income distribution is the Lorenz curve. Income inequality is usually measured by the area enclosed by the Lorenz curves for the income distribution and perfect equality. The area results in the Gini coefficient. This section attempts to derive an income inequality index from the Lorenz curve for the UD income distribution. Similar to the Lorenz curve for income, the Lorenz curve for the UD income is defined as

$$L(p) = \frac{\int_0^p Q(t)dt}{\mu_x}.$$

However, the Lorenz curve for UD income is not applicable, because the Lorenz curve for $\mathbf{x}_{pe}$ is not defined due to $\mu_x = 0$. Thus, we use $\int_0^p Q(t)dt$, which is referred to as the unscaled Lorenz curve and denoted by $L^u(p)$.

Figure 2 depicts the unscaled Lorenz curves $L^u(p)$ and $L_{pe}^u(p)$ for $\mathbf{x}$ and $\mathbf{x}_{pe}$. Note that $L_{pe}^u(p)$ lies on the horizontal axis. The departure of $L^u(p)$ from $L_{pe}^u(p)$ can be assessed by area **B**, which is the ℓ_1 norm of the departure. The dotted diagonal line is not an unscaled Lorenz curve and does not represent perfect equality. The dotted diagonal line is a reference line for computing area **A** enclosed by $L^u(p)$ and the dotted line. Since area **A** is $\mu_x G_x/2$, area **B** is obtained as

$$\mathbf{B} = \frac{1}{2} \mu_x (1 - G_x).$$

Normalizing **B** by μ_y , we have

$$\frac{\mathbf{B}}{\mu_y} = \frac{1}{2} \frac{\mu_x}{\mu_y} (1 - G_x) = \frac{1}{2} (\tilde{\ell}_2^v)^2. \quad (6)$$

Fig. 2 Unscaled Lorenz curves $L^u(p)$ and $L_{pe}^u(p)$ for $\mathbf{x}$ and $\mathbf{x}_{pe}$

The normalized ℓ_1 norm of the difference between $L^u(p)$ and $L_{pe}^u(p)$ is equivalent to $\tilde{\ell}_2^v$ obtained in the previous section. That is, the CDF and the unscaled Lorenz curve for UD income result in equivalent indexes.

5 Application of normalized norm indexes

In this section, we apply the normalized norm indexes to example income distributions and real income datasets. First, consider the eight example income distributions in Table 1. The first seven income distributions $\mathbf{y}_i, i = 1, 2, \dots, 7$ have the same $n = 5$, $S_y = 15$, and $\mu_y = 3$. Each of them is obtained from others by a series of transfers. So we can investigate the effect of transfers on income inequality. There are two couples of income distributions that G_y and CV_y fail to distinguish with respect to income inequality, but the normalized norm indexes succeed. In order to examine the negative income problem, we also included an income distribution with a negative income.

G_y and CV_y assess the income inequality of $\mathbf{y}_1$ and $\mathbf{y}_2$ as being the same. However, there is a difference between $\mathbf{y}_1$ and $\mathbf{y}_2$; 5 of S_y is evenly distributed among the individuals in $\mathbf{y}_1$, while none of S_y is evenly distributed among the individuals in $\mathbf{y}_2$. This difference is revealed by the Rawlsian index $\tilde{\ell}_1$. $\mathbf{y}_1$ has less $\tilde{\ell}_1$ than does $\mathbf{y}_2$. Although $\mathbf{y}_2$ has less G_x and CV_x than $\mathbf{y}_1$, $\mathbf{y}_1$ has less $\tilde{\ell}_2^v$ and $\tilde{\ell}_2^h$ than $\mathbf{y}_2$ because it has less $\tilde{\ell}_1$. Unlike G_y and CV_y , the normalized norm indexes indicate that $\mathbf{y}_1$ is more equal than $\mathbf{y}_2$. A similar result is obtained from a comparison of $\mathbf{y}_3$ and $\mathbf{y}_4$. We can consider a series of progressive transfers preserving the mean UD income that transform $\mathbf{y}_1$ into $\mathbf{y}_4$. According to G_y and CV_y , $\mathbf{y}_1$ is more equal than $\mathbf{y}_4$. While $\mathbf{y}_1$ and $\mathbf{y}_4$ have the same Rawlsian index, $\mathbf{y}_4$ has less G_x and CV_x than $\mathbf{y}_1$. However, according to $\tilde{\ell}_2^v$ and $\tilde{\ell}_2^h$, $\mathbf{y}_4$ is more vertically unequal and horizontally equal than $\mathbf{y}_1$. A progressive transfer preserving the mean UD income decreases horizontal inequality but increases the vertical inequality. That is, a progressive transfer preserving the mean UD income does not reduce income inequality. We can observe this phenomenon when

Table 1 Example income distributions

y	G_y	CV_y	$\tilde{\ell}_1$	G_x	CV_x	$\tilde{\ell}_2^v$	$\tilde{\ell}_2^h$
$y_1 = (1, 2, 2, 5, 5)$	0.2933	0.5578	0.6667	0.4400	0.8367	0.6110	0.8692
$y_2 = (0, 3, 3, 4, 5)$	0.2933	0.5578	1.0000	0.2933	0.5578	0.8406	1.1450
$y_3 = (2, 2, 2, 4, 5)$	0.2133	0.4216	0.3333	0.6400	1.2649	0.3464	0.5375
$y_4 = (1, 3, 3, 3, 5)$	0.2133	0.4216	0.6667	0.3200	0.6325	0.6733	0.7888
$y_5 = (1, 3, 3, 4, 4)$	0.1867	0.3651	0.6667	0.2800	0.5477	0.6928	0.7601
$y_6 = (2, 3, 3, 3, 4)$	0.1067	0.2108	0.3333	0.3200	0.6325	0.4761	0.3944
$y_7 = (-1, 4, 4, 4, 4)$	0.2667	0.6667	1.3333	0.2000	0.5000	1.0328	1.4907
$y_8 = (0, 4, 4, 4, 4)$	0.2000	0.5000	1.0000	0.2000	0.5000	0.8944	1.1180

we compare y_4 and y_5 . y_6 is obtained from y_5 by a progressive transfer reducing the mean UD income, i.e., the transfer reducing the Rawlsian index. Such progressive transfers decrease both horizontal income inequality and vertical income inequality and consequently decrease income inequality.

Let us consider y_7 in which a negative income value is observed. G_y and CV_y for y_7 in Table 1 are computed without any manipulation of the negative income value. If the negative income value is removed from y_7 , y_7 will exhibit perfect equality. If the negative income value is replaced with zero, y_7 will become y_8 . Replacement with zero primarily reduces UD income and the Rawlsian index. Consequently, income inequality is underestimated as seen from $\tilde{\ell}_1$ s, $\tilde{\ell}_2^v$ s and $\tilde{\ell}_2^h$ s for y_7 and y_8 . According to G_y , y_7 is more unequal than y_1 and y_2 . However, the normalized norm indexes assume that y_7 is the most unequal income distribution among the income distributions in Table 1.

Next, we apply the normalized norm indexes to the latest LIS datasets for household disposable income after the year 2010 (LIS 2019). Because the years included in the datasets are not the same, this analysis is not intended to be a thorough cross-national comparison but is instead intended to test the applicability of the normalized norm indexes to cross-national comparison. The normalized norm indexes and the popular Gini coefficients and CVs for forty-two countries were computed and are presented in Table 2. The countries were arranged in increasing order of G_y . The numbers in parentheses are ranks for each index. G_y 's in Table 2 are slightly different from the Gini coefficients published by the LIS Data Center. This is because the LIS Data Center computes the Gini coefficient after removing negative and zero income values.

We first note that negative incomes are frequently observed. The Rawlsian index is greater than one when there are negative incomes. Twenty-seven countries have negative incomes. The normalized norm indexes provide considerably different information on income inequality than G_y and CV_y . The Gini coefficient indicates that Denmark and Norway belong to five most equal countries. However, all the normalized norm indexes indicate that Denmark and Norway are among the five most unequal countries. According to G_y and CV_y , Paraguay and South Africa are among the six most unequal countries. However, Paraguay and South Africa are two most equal countries from the perspective of $\tilde{\ell}_2^h$.

Table 2 Analysis of the latest LIS datasets after the year 2010

Country	Year	G_y	CV_y	$\tilde{\ell}_1$	$\tilde{\ell}_2^v$	$\tilde{\ell}_2^h$
Iceland	2010	0.2453(1)	0.5767(3)	1.0000(8)	0.8687(17)	1.1544(4)
Denmark	2013	0.2539(2)	0.6529(14)	22.1043(39)	4.6744(39)	22.1139(39)
Norway	2013	0.2540(3)	0.6381(12)	146.1945(42)	12.0806(42)	146.1959(42)
Czech	2013	0.2587(4)	0.5838(5)	0.8762(1)	0.7858(7)	1.0529(1)
Finland	2013	0.2611(5)	0.6280(10)	1.0051(16)	0.8626(16)	1.1851(6)
Netherlands	2013	0.2672(6)	0.5921(7)	2.3173(28)	1.4318(28)	2.3917(27)
Hungary	2015	0.2682(7)	0.5557(2)	0.9486(2)	0.8248(12)	1.0994(2)
Slovakia	2013	0.2696(8)	0.5817(4)	1.0000(9)	0.8546(15)	1.1569(5)
Slovenia	2012	0.2711(9)	0.5509(1)	0.9911(3)	0.8485(14)	1.1339(3)
Austria	2013	0.2807(10)	0.6593(15)	1.0000(9)	0.8481(13)	1.1978(7)
Luxembourg	2013	0.2874(11)	0.5974(9)	2.7989(32)	1.5848(32)	2.8619(32)
France	2010	0.2922(12)	0.7415(26)	2.4573(29)	1.4714(29)	2.5668(28)
Switzerland	2013	0.2962(13)	0.6763(18)	1.1703(19)	0.9349(20)	1.3517(14)
Ireland	2010	0.2963(14)	0.5899(6)	1.9036(25)	1.2678(25)	1.9929(23)
Germany	2015	0.2988(15)	0.7845(27)	7.0353(34)	2.5955(34)	7.0789(34)
Poland	2016	0.3006(16)	0.6750(17)	23.9225(40)	4.8602(40)	23.9321(40)
Taiwan	2016	0.3035(17)	0.6369(11)	1.1852(20)	0.9390(21)	1.3455(13)
South Korea	2012	0.3071(18)	0.5931(8)	1.2601(23)	0.9762(23)	1.3927(15)
Canada	2013	0.3215(19)	0.6413(13)	2.2300(27)	1.3815(27)	2.3203(26)
Russia	2016	0.3315(20)	0.6901(21)	0.9986(5)	0.8167(11)	1.2139(8)
Australia	2014	0.3323(21)	1.0590(33)	2.5526(30)	1.4901(30)	2.7636(30)
Italy	2014	0.3326(22)	0.7130(23)	1.1280(17)	0.8919(18)	1.3344(12)
Greece	2013	0.3347(23)	0.7402(25)	1.2348(22)	0.9487(22)	1.4396(16)
Serbia	2016	0.3348(24)	0.6815(19)	7.5408(36)	2.6844(36)	7.5715(36)
Spain	2013	0.3468(25)	0.6817(20)	2.1015(26)	1.3247(26)	2.2093(25)
Israel	2016	0.3476(26)	0.6692(16)	5.4256(33)	2.2534(33)	5.4667(33)
Lithuania	2013	0.3489(27)	0.7202(24)	1.0000(9)	0.8069(10)	1.2323(9)
UK	2016	0.3517(28)	0.9258(31)	11.9430(38)	3.4046(38)	11.9788(38)
Estonia	2013	0.3553(29)	0.7017(22)	2.6384(31)	1.5110(31)	2.7302(29)
Uruguay	2016	0.3604(30)	0.8084(28)	1.0000(9)	0.7998(9)	1.2859(10)
US	2016	0.3862(31)	0.8745(30)	1.2274(21)	0.9172(19)	1.5071(18)
Georgia	2016	0.3903(32)	0.8292(29)	1.0000(9)	0.7808(6)	1.2991(11)
Guatemala	2014	0.4160(33)	1.5343(39)	1.0000(9)	0.7642(5)	1.8314(21)
Brazil	2013	0.4601(34)	1.2164(35)	0.9979(4)	0.7333(4)	1.5734(19)
Peru	2013	0.4641(35)	1.0483(32)	29.5987(41)	5.3976(41)	29.6173(41)
Panama	2013	0.4721(36)	1.1166(34)	1.0000(9)	0.7266(3)	1.4990(17)
Mexico	2012	0.4834(37)	1.2853(36)	7.3350(35)	2.6176(35)	7.4468(35)
Paraguay	2016	0.4885(38)	1.4582(37)	0.9988(6)	0.7144(2)	1.7675(20)

Table 2 continued

Country	Year	G_y	CV_y	$\tilde{\ell}_1$	$\tilde{\ell}_2^v$	$\tilde{\ell}_2^h$
India	2011	0.5069(39)	1.5731(40)	10.0005(37)	3.0812(37)	10.1235(37)
Colombia	2013	0.5158(40)	1.4821(38)	1.1371(18)	0.7882(8)	1.8680(22)
Egypt	2012	0.5202(41)	2.2471(42)	1.7323(24)	1.1010(24)	2.8373(31)
South Africa	2012	0.6073(42)	1.7890(41)	0.9999(7)	0.6265(1)	2.0495(24)

Fig. 3 Plot of $\tilde{\ell}_2^v$ against $\tilde{\ell}_2^h$

Less vertical departure does not mean less horizontal departure. That is, less $\tilde{\ell}_2^v$ is not always accompanied by less $\tilde{\ell}_2^h$. The ranks for $\tilde{\ell}_2^v$ are not the same as the ranks for $\tilde{\ell}_2^h$. $\tilde{\ell}_2^v$ is plotted against $\tilde{\ell}_2^h$ in Fig. 3. Figure 3 shows that less $\tilde{\ell}_2^v$ can be accompanied by greater $\tilde{\ell}_2^h$, especially when both $\tilde{\ell}_2^v$ and $\tilde{\ell}_2^h$ are small.

The countries with the eleven largest Rawlsian indexes are Norway, Peru, Poland, Denmark, the UK, India, Serbia, Mexico, Germany, Israel, and Luxembourg. These countries correspond to the countries with the eleven largest $\tilde{\ell}_2^v$ and $\tilde{\ell}_2^h$. The ranks of these countries on the Rawlsian index, $\tilde{\ell}_2^v$ and $\tilde{\ell}_2^h$ are the same. The countries with the five smallest Rawlsian indexes are the Czech Republic, Hungary, Slovenia, Brazil, and Russia. Their ranks for $\tilde{\ell}_2^v$ and $\tilde{\ell}_2^h$ are low but somewhat different from one another. This indicates that the Rawlsian index dominates $\tilde{\ell}_2^v$ and $\tilde{\ell}_2^h$ when the Rawlsian index is large. The $\tilde{\ell}_2^v$'s and $\tilde{\ell}_2^h$'s are plotted against $\tilde{\ell}_1$ in Figs. 4 and 5. Figures 4 and 5 show that there is a strong linear tendency between the Rawlsian index and the normalized ℓ_2 norm indexes. Exceptional cases from this linear tendency are observed when the Rawlsian index is small. When the Rawlsian index is small, the influences of G_x and CV_x , can surpass that of the Rawlsian index.

Fig. 4 Plot of $\tilde{\ell}_2^v$ against $\tilde{\ell}_1$

Fig. 5 Plot of $\tilde{\ell}_2^h$ against $\tilde{\ell}_1$

For a given income distribution, a progressive transfer maintaining the Rawlsian index decreases $\tilde{\ell}_2^h$ but increases $\tilde{\ell}_2^v$. Although the income distributions in Table 2 are different, we can observe similar negative correlations between $\tilde{\ell}_2^h$ and $\tilde{\ell}_2^v$ when the Rawlsian index is fixed. Figure 6 is a plot of $\tilde{\ell}_2^v$ against $\tilde{\ell}_2^h$ for seven countries with $\tilde{\ell}_1 = 1$. The plot shows that $\tilde{\ell}_2^h$ and $\tilde{\ell}_2^v$ are negatively correlated.

6 Conclusions

Measurement is the foundation of quantitative research. The proper measurement of income inequality is important for studying economic inequality. In this paper,

Fig. 6 Plot of $\tilde{\ell}_2^v$ against $\tilde{\ell}_2^h$ for the countries with $\tilde{\ell}_1 = 1$

we proposed a new framework for measuring income inequality. The framework is based on the UD income distribution, which is obtained from the income distribution by removing the equally distributed parts of incomes. We then derived the normalized norm indexes from the CDF and the unscaled Lorenz curve for the UD income distribution. The normalized norm indexes represent the location and the dispersion of the UD income distribution, while most conventional income inequality indexes reflect only the dispersion of the income distribution. By analyzing example income distributions and the LIS datasets, we showed that the normalized norm indexes are appropriate measures of income inequality and solve the negative income problem. However, it should be noted that the proposed indexes represent only one-dimensional views of income inequality. They describe the horizontal and vertical dimensions of the two-dimensional UD income distribution. Therefore, the horizontal and vertical normalized norm indexes are not alternatives to one another. Both the horizontal and vertical normalized norm indexes should be taken into consideration to comprehensively understand income inequality. For the sake of simplicity without sacrificing comprehensiveness, it would be desirable to devise a new index that reflects a two-dimensional view of income inequality. Our future research will be directed toward the development of such an index.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Atkinson AB (1970) On the measurement of inequality. *J Econ Theory* 2:244–263
- Bellù LG, Liberati P (2006) Describing income inequality: Theil index and entropy class indices. http://www.fao.org/docs/up/easypol/445/theil_index_051en.pdf
- Chen CN, Tsau TW, Rhab TS (1982) The Gini coefficient and negative income. *Oxford Econ Pap* 34:473–478
- Clementi F, Gallegati M, Gianmoena L, Landini S, Stiglitz JE (2019) Mis-measurement of inequality: a critical reflection and new insights. *J Econ Interac Coord* 14:891–921
- Cowell FA (2011) *Measuring inequality*, 3rd edn. Oxford University Press, New York
- Donaldson D, Weymark JA (1980) A single-parameter generalization of the Gini indices of inequality. *J Econ Theory* 22:67–86
- Dorfman R (1979) A Formula for the Gini coefficient. *Rev Econ Stat* 16(10):146–149
- Gallegati M, Landini S, Stiglitz JE (2016) The inequality multiplier. Research paper 16–29. Columbia Business School, New York
- Gastwirth JL (2014) Median-based measures of inequality: reassessing the increase in income inequality in the U.S. and Sweden. *J IAOS* 30:311–20
- Gini C (1914) Sulla misura della concentrazione e della variabilità dei caratteri. *Atti del Reale Istituto veneto di scienze, lettere ed arti* 73:1201–1248 [english translation: On the measurement of concentration and variability of characters, *Metron*, 63:3–38, 2005]
- Hao L, Naiman DQ (2010) *Assessing inequality*. Sage Publications Inc, New York
- Jenkins SP, Kerm PV (2011) The measurement of economic inequality. In: Nolan B, Salverda W, Smeeding TM (eds) *The Oxford handbook of economic inequality*. Oxford University Press, New York, pp 40–67
- Lorenz O (1905) Methods of measuring the concentration of wealth. *Publ Am Stat Assoc* 9:209–2019
- Luxembourg Income Study (LIS) Database. LIS, Luxembourg. <http://www.lisdatacenter.org>. Accessed 11 May 2019
- OECD (2013) OECD framework for statistics on the distribution of household income, consumption, and wealth. <http://doi.org/10.1787/9789264194830-en>. Accessed 28 Mar 2016
- Park JY, Kim YS, Heo SJ (2018) Dual-index measurement of income inequality. *Bull Econ Res* 70(3):277–284
- Raffinetti E, Siletti E, Vernizzi A (2015) On the Gini coefficient normalization when incomes with negative values are considered. *Stat Methods Appl* 24(3):507–521
- Raffinetti E, Siletti E, Vernizzi A (2016) Analyzing the effects of negative and non-negative values on income inequality: evidence from the survey of household income and Wealth of the Bank of Italy (2012). *Soc Indic Res* 133:185–207
- Schutz RR (1951) On the measurement of income inequality. *Am Econ Rev* 41:107–122
- Shorrocks AF (1980) The class of additively decomposable inequality measures. *Econometrica* 48:613–625
- Theil H (1967) *Economics and information theory*. North Holland, Amsterdam
- Yitzhaki S (1983) On an extension of the Gini inequality index. *Int Econ Rev* 24:617–628
- Yitzhaki S (1998) More than a dozen alternative ways of spelling Gini. *Res Econ Inequal* 8:13–30
- Zanardi G (1964) Della asimmetria condizionata delle curve di concentrazione. Lo scontentamento. *Rivista Italiana di Economia, Demografia e Statistica* 18:431–466
- Zenga M (2007) Inequality curve and inequality index based on the ratios between lower and upper arithmetic means. *Stat Appl* 5:3–27

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.