ECINEQ
Society for the Study of Economic Inequality

Working Paper Series

A Paradox for Inequality Indices

Satya Chakravarty
Palash Sarkar

ECINEQ 2020 559

ECINEQ 2020 559
2020 August
www.ecineq.org

A Paradox for Inequality Indices

Satya Chakravarty
Indian Statistical Institute
Palash Sarkar
Indian Statistical Institute

Abstract

This paper identifies a paradox for inequality indices which is similar to the well known Simpson's paradox in statistics. For the Gini and Bonferroni indices, concrete examples of the paradox are provided and general methods are described for obtaining examples of the paradox for arbitrary sizepopulation.

Keyword: inequality, paradox, Gini index, Bonferroni index, Simpson's paradox

JEL Cassification: C43, D31, D63, O15

A Paradox for Inequality Indices

Satya R. Chakravarty and Palash Sarkar
Indian Statistical Institute
203, B.T. Road
Kolkata
India 700108
email:{satya, palash}@isical.ac.in

August 11, 2020

Abstract

This paper identifies a paradox for inequality indices which is similar to the well known Simpson's paradox in statistics. For the Gini and Bonferroni indices, concrete examples of the paradox are provided and general methods are described for obtaining examples of the paradox for arbitrary size population.

Key Words: inequality, paradox, Gini index, Bonferroni index, Simpson's paradox.

JEL Classification Codes: C43, D31, D63, O15.

1 Introduction

Understanding income inequality in a society is of paramount importance in social planning and development. A basic requirement for understanding the issue is the ability to measure inequality. There is a rich literature on various indices for measuring inequality exploring both theoretical properties as well as applications in policy planning. We refer to Cowell (2016) for a recent survey on inequality measures.

A purpose of inequality measurement is to be able to compare inequalities of societies or across time. Income, on the other hand, is often an aggregated value. For example, an individual may have more than one sources of income and the total income of the individual is the sum of the individual incomes. Another example would be the income of a family is the aggregate of the incomes of the individuals in the family. Yet another example is that income over a number of years is the aggregate of the incomes in the individual years. In such cases, one would often measure and compare inequalities in the components of the income as well as measure and compare inequality in the aggregate income.

In this paper, we demonstrate a paradox that can arise when one is measuring and comparing inequality in component as well as aggregated incomes. Suppose there are two sets of people, where the income of any individual in any of the sets has two components which is aggregated to obtain the total income of the individual. We show that the following paradoxical situation may arise. For the first set of people, the inequalities in the income distributions of both the components is greater than the corresponding inequalities for the second set of people, yet the inequality in the aggregate income for the first set of people is less than that of the second set. A formal description and suggestive examples are provided later.

The paradox that we describe is similar to the well known Simpson paradox in statistics. This paradox refers to the phenomenon where a clear direction emerges in individual data sets, but disappears or reverses in the aggregated data. The effect was identified by Simpson (1951), but had been mentioned earlier by Pearson et al. (1899) and Yule (1903). While these three papers had noticed the disappearance of sub-population trend in aggregated population, the reversal phenomenon was noted by Choen and Nagel (1934). The peculiarity of the reversal and its difficulty in interpretation was termed a paradox by Blyth (1972).

We have chosen the Gini index to demonstrate the existence of the above described paradoxical feature in inequality measurement since it is the most widely used index of inequality in applied works (see Sen (1973), Donaldson-Weymark (1980))[1]. This could be due to many attractive properties including its geometric interpretation with respect to the well known Lorenz curve. It is a normalized sum of pair-wise absolute income differences and hence is easy to understand and compute. In contrast to many other inequality metrics, it can easily accommodate non-positive incomes. If the rank-order of incomes is the same for all sources of income, the Gini index for the aggregate lncome distribution is simply the mean-weighted sum of the source-wise Gini indices, where the weights are normalized to sum to unity (Weymark (1981)). The second inequality metric we choose for demonstrating the paradox is the Bonferroni index (Bonferroni (1930)), which is unambiguously bounded from below by the Gini index (Chakravarty and Sarkar (2020)). For a non-decreasingly ordered income distribution, it is a normalized average of differences between the overall mean and the partial means. It can as well accommodate non-positive incomes. It has a geometric interpretation in terms of the Bonferroni curve, a plot of the ratio between the cumulative income shares and cumulative population proportions against cumulative population proportions (Aaberge (2007) and Barcena-Martin and Silber (2013)).

Concrete examples of the paradox are provided for the Gini and the Bonferroni indices. More generally, starting from an instance of the Simpson's paradox, we describe methods for obtaining examples

[1] The World Bank site `https://www.worldbank.org/en/topic/poverty/lac-equity-lab1/income-inequality/inequality-trends` mentions the Gini index as one of the most widely used inequality indicators.

of the paradox for arbitrary size population for both these inequality indices. As discussed above, income is often obtained as aggregate of component incomes. The paradox raises questions regarding the interpretation of comparing inequalities of two populations whose incomes are obtained by aggregating component incomes.

2 The Paradox

Let I be an inequality index on a set of income distributions. We identify a possible paradoxical condition for I.

Paradox:

Suppose there are income distributions $\mathbf{x}$, $\mathbf{x}'$, $\mathbf{y}$ and $\mathbf{y}'$ such that the following condition holds.

$$I(\mathbf{x}) > I(\mathbf{x}') \text{ and } I(\mathbf{y}) > I(\mathbf{y}'), \quad \textit{but} \quad I(\mathbf{x}+\mathbf{y}) \leq I(\mathbf{x}'+\mathbf{y}'). \tag{1}$$

Then the paradox holds for I.

We illustrate the above through some examples.

Example 1: Suppose that the individuals in a society have an income and also obtain subsidies from the government. Let $\mathbf{x}$ and $\mathbf{x}'$ be the income distributions of the individuals in the society in two successive years. Further, let $\mathbf{y}$ and $\mathbf{y}'$ be the subsidies received by the individuals in two successive years. The combined income and subisides in the two successive years are $\mathbf{x}+\mathbf{y}$ and $\mathbf{x}'+\mathbf{y}'$. Suppose that (1) holds. The inequalities $I(\mathbf{x}) > I(\mathbf{x}')$ and $I(\mathbf{y}) > I(\mathbf{y}')$ mean that individually, the inequalities in both incomes and subsidies in the first year are more than that in the second year. On the other hand, the condition $I(\mathbf{x}+\mathbf{y}) < I(\mathbf{x}'+\mathbf{y}')$ means that when one considers the combined income-plus-subsidy, the inequality actually goes down in the second year compared to the first year. So, the paradox here is that even though individually the inequalities in the income and subsidies are higher in first year compared to the second year, the combined distribution of income-plus-subsidy has lower inequality in the first year compared to the second year. We can as well interpret $\mathbf{x}$ and $\mathbf{x}'$ as distributions of incomes earned by individuals in a year from two different sources. An analogous interpretation holds for $\mathbf{y}$ and $\mathbf{y}'$. Then the paradox refers to a situation in which we are concerned with inequality rankings of aggregate income distributions and their components.

Example 2: Suppose that the individuals in a society are divided into two social groups (say, male and female). Further, suppose that both the groups have the same number of individuals. Let $\mathbf{x}$ and $\mathbf{y}$ be the income distributions of males in two successive years. Similarly, let $\mathbf{x}'$ and $\mathbf{y}'$ be the income distributions of females in two successive years. Considering both the years together, the income distributions of the male and female groups are $\mathbf{x}+\mathbf{y}$ and $\mathbf{x}'+\mathbf{y}'$ respectively. Suppose that (1) holds. The conditions $I(\mathbf{x}) > I(\mathbf{x}')$ and $I(\mathbf{y}) > I(\mathbf{y}')$ mean that individually, in both the years the inequalities in the income distributions of males are more than the inequalities in the income distributions of females. On the other hand, the condition $I(\mathbf{x}+\mathbf{y}) < I(\mathbf{x}'+\mathbf{y}')$ means that considering a two-year period, the inequality in the income distribution of males is actually lower than the inequality in the income distribution of females. So, the paradox here is that considering one year at a time, income distributions of males have higher inequalities than that of females, whereas considering a two-year period, the income distribution of males is less unequal than that of females.

Example 3: Suppose that the individuals in a society are divided into two social groups $\mathcal{G}_1$ and $\mathcal{G}_2$. Further, suppose both the groups have the same number of families with each family having a single male earning member and a single female earning member. Let $\mathbf{x}$ and $\mathbf{y}$ be the income distributions of males females respectively of families in $\mathcal{G}_1$. Similarly, let $\mathbf{x}'$ and $\mathbf{y}'$ be the income distributions of males and females respectively of families in $\mathcal{G}_2$. Then the income distribution of the families in $\mathcal{G}_1$ is $\mathbf{x}+\mathbf{y}$, while the income distribution of families in $\mathcal{G}_2$ is $\mathbf{x}'+\mathbf{y}'$. Suppose that (1) holds. The conditions $I(\mathbf{x}) > I(\mathbf{x}')$ and $I(\mathbf{y}) > I(\mathbf{y}')$ mean that individually the inequalities in the income distributions of males and females in $\mathcal{G}_1$ are more than the corresponding inequalities in the income distributions of males and females in $\mathcal{G}_2$. On the other hand, the condition $I(\mathbf{x}+\mathbf{y}) < I(\mathbf{x}'+\mathbf{y}')$ means that considering income distributions of families, the inequality in the income distribution of families in $\mathcal{G}_1$ is less than the inequality in the income distribution of families in $\mathcal{G}_2$. So, the paradox here is that individually considering males and females, the inequalities in income distributions in $\mathcal{G}_1$ are higher than those of $\mathcal{G}_2$, but considering families, the inequality in the income distribution in $\mathcal{G}_1$ is lower than that in $\mathcal{G}_2$.

Dual form of the paradox: The inequality reversal in (1) is from 'greater than' to 'less than'. A dual form of the paradox can be stated where the inequality reversal is from 'less than' to 'greater than'.

Suppose there are income distributions $\mathbf{x}$, $\mathbf{x}'$, $\mathbf{y}$ and $\mathbf{y}'$ such that the following condition holds.

$$I(\mathbf{x}) < I(\mathbf{x}') \text{ and } I(\mathbf{y}) < I(\mathbf{y}'), \quad \textit{but} \quad I(\mathbf{x}+\mathbf{y}) \geq I(\mathbf{x}'+\mathbf{y}'). \tag{2}$$

Then the dual form of the paradox holds for I.

3 Concrete Examples of the Paradox

We provide concrete examples of the paradox for two well known inequality indices, namely, the Gini and the Bonferroni indices. Before describing the examples, we provide the formal definitions of these inequality indices.

Fix an integer $n > 2$. An income distribution in a society with n individuals is a vector $\mathbf{x} = (x_1, \ldots, x_n)$, where $x_1 \leq x_2 \leq \cdots \leq x_n$ and $x_1 + \cdots + x_n > 0$. Let D_n be the set of all income distributions in a society with n individuals. An inequality index I is a real valued function defined on the set of income distributions. Formally, $I : D_n \to \mathbb{R}$, where $\mathbb{R}$ is the set of real numbers.

Let $\mathbf{x} \in D_n$ be arbitrary. For $i = 1, \ldots, n$, define $s_i = x_1 + \cdots + x_i$ and $\mu_i = s_i/i$. Given $\mathbf{x} \in D_n$, the Gini index of $\mathbf{x}$ is defined to be $G(\mathbf{x})$, where

$$G(\mathbf{x}) \quad = \quad \frac{1}{2n^2\mu_n}\sum_{i=1}^{n}\sum_{j=1}^{n}|x_i - x_j|. \tag{3}$$

Given $\mathbf{x} \in D_n$, the Bonferroni index of $\mathbf{x}$ is defined to be $B(\mathbf{x})$, where $B(\mathbf{x})$ is given by the following expression.

$$B(\mathbf{x}) \quad = \quad \frac{1}{\mu_n}\left(\mu_n - \frac{1}{n}\sum_{i=1}^{n}\mu_i\right) = \frac{1}{n\mu_n}\left(\sum_{i=1}^{n}(x_i - \mu_i)\right). \tag{4}$$

For both the Gini and the Bonferroni indices we provide examples of the paradox in its original and dual form.

Paradox for the Gini index: Let $\mathbf{x} = (2, 7, 7)$, $\mathbf{x}' = (104, 182, 234)$, $\mathbf{y} = (10, 12, 18)$ and $\mathbf{y}' = (65, 80, 117)$. Then

$$G(\mathbf{x}) = 0.208333 > 0.166667 = G(\mathbf{x}') \text{ and } G(\mathbf{y}) = 0.133333 > 0.132316 = G(\mathbf{y}'), \textit{ but}$$
$$G(\mathbf{x}+\mathbf{y}) = 0.154762 < 0.155158 = G(\mathbf{x}'+\mathbf{y}').$$

So, (1) holds.

Let $\mathbf{x} = (8, 24, 32)$, $\mathbf{x}' = (1, 4, 5)$, $\mathbf{y} = (1, 2, 3)$ and $\mathbf{y}' = (16, 20, 44)$.

$$G(\mathbf{x}) = 0.25 < 0.266667 = G(\mathbf{x}') \text{ and } G(\mathbf{y}) = 0.222222 < 0.233333 = G(\mathbf{y}'), \textit{ but}$$
$$G(\mathbf{x}+\mathbf{y}) = 0.247619 > 0.237037 = G(\mathbf{x}'+\mathbf{y}').$$

So, (2) holds.

Paradox for the Bonferroni index: Let $\mathbf{x} = (6, 16, 26)$, $\mathbf{x}' = (240, 580, 740)$, $\mathbf{y} = (24, 40, 56)$ and $\mathbf{y}' = (75, 143, 175)$. Then

$$B(\mathbf{x}) = 0.3125 > 0.25 = B(\mathbf{x}') \text{ and } B(\mathbf{y}) = 0.2 > 0.198473 = B(\mathbf{y}'), \textit{ but}$$
$$B(\mathbf{x}+\mathbf{y}) = 0.232143 < 0.239631 = B(\mathbf{x}'+\mathbf{y}').$$

So, (1) holds.

Let $\mathbf{x} = (6, 38, 52)$, $\mathbf{x}' = (6, 22, 47)$, $\mathbf{y} = (1, 3, 5)$ and $\mathbf{y}' = (45, 245, 310)$.

$$B(\mathbf{x}) = 0.375 < 0.4 = B(\mathbf{x}') \text{ and } B(\mathbf{y}) = 0.333333 < 0.35 = B(\mathbf{y}'), \textit{ but}$$
$$B(\mathbf{x}+\mathbf{y}) = 0.371429 > 0.35556 = B(\mathbf{x}'+\mathbf{y}').$$

So, (2) holds.

4 Simpson's Paradox

As mentioned earlier, the paradox we have formulated in (1) (and 2) has similarities with Simpson's paradox. We provide a simple example of Simpson's paradox from Grimmett and Stirzaker (2001). Suppose a clinical trial is being performed to compare the efficacies of two drugs, say Drug-I and Drug-II. Both the drugs are administered to males and females. Success and failures are individually measured. Suppose Drug-I is given to 2000 females and 20 males; 200 of the females are successful, while 19 of the males are successful. Further, suppose Drug-II is given to 200 females and 2000 males; 10 of the females are successful, while 1000 of the males are successful. Considering females and males separately, the success rates of Drug-I for females and males are 1/10 and 19/20 which are respectively greater than 1/20 and 1/2, the corresponding success rates of Drug-II for females and males. So, considered individually among females and males, Drug-I is better than Drug-II. Now consider the overall success rate, where gender is ignored. Drug-I has been administered to 2020 persons, out of which 219 are successful, while Drug-II has been administered to 2200 persons of which 1010 are successful. So, considering the overall picture, Drug-I has success rate 219/2020 which is less than 1010/2200, the success rate of Drug-II. The paradox is that while in the individual groups Drug-I is seen to be more efficacious, in the combined group Drug-II becomes more efficacious.

More generally, Simpson's paradox can be seen to arise if there are positive numbers $a_1, r_1, a_2, r_2, b_1, t_1, b_2, t_2$ with $a_1 < r_1$, $a_2 < r_2$, $b_1 < t_1$ and $b_2 < t_2$ such that

$$\frac{a_1}{r_1} > \frac{b_1}{t_1} \text{ and } \frac{a_2}{r_2} > \frac{b_2}{t_2} \textit{ but } \frac{a_1 + a_2}{r_1 + r_2} < \frac{b_1 + b_2}{t_1 + t_2}. \tag{5}$$

Here the inequality reversal is of from 'greater than' to 'less than'. This corresponds to the type of paradox given in (1). One can also formulate a version of Simpson's paradox where inequality reversal is of the type 'less than' to 'greater than' which would correspond to the type of the paradox given in (2).

We have provided one concrete example of Simpson's paradox. Other examples are known and have been reported in the literature. See for example Bandyopadhyay et al. (2010).

5 Obtaining Examples of the Inequality Paradox

We have provided numerical examples of the inequality paradox for both the Gini and Bonferroni indices. These are for $n = 3$. In this section, we briefly describe a method for obtaining examples of the paradox for general values of n.

First consider the Gini index. Since we are considering the income distribution $\mathbf{x}$ to be non-decreasingly ordered, the definition of the Gini index given in (3) can be equivalently written as

$$G(\mathbf{x}) = \frac{1}{ns_n} \sum_{1\leq i<j\leq n} (x_j - x_i) = \frac{1}{n} \cdot \frac{\sum_{1\leq i<j\leq n}(x_j - x_i)}{x_1 + \cdots + x_n} = \frac{1}{n} \cdot \frac{\alpha_1 x_1 + \cdots + \alpha_n x_n}{x_1 + \cdots + x_n} \tag{6}$$

where $\alpha_i = 2i - 1 - n$, for $i = 1, \ldots, n$. For a fixed n, $\alpha_1, \ldots, \alpha_n$ are constants.

We use (6) for obtaining an example of the inequality paradox for the Gini index. We start with an example of Simpson's paradox. Suppose a_1, r_1, a_2, r_2 and b_1, t_1, b_2, t_2 are such that (5) holds. Further, suppose we can find income distributions $\mathbf{x} = (x_1, \ldots, x_n)$, $\mathbf{y} = (y_1, \ldots, y_n)$, $\mathbf{x}' = (x'_1, \ldots, y'_n)$ and $\mathbf{y}' = (y_1, \ldots, y'_n)$ such that

$$\left.\begin{array}{rr} \alpha_1 x_1 + \cdots + \alpha_n x_n = a_1; & x_1 + \cdots + x_n = r_1; \\ \alpha_1 y_1 + \cdots + \alpha_n y_n = a_2; & y_1 + \cdots + y_n = r_2; \\ \alpha_1 x'_1 + \cdots + \alpha_n x'_n = b_1; & x'_1 + \cdots + x'_n = t_1; \\ \alpha_1 y'_1 + \cdots + \alpha_n y'_n = b_2; & y'_1 + \cdots + y'_n = t_2. \end{array}\right\} \tag{7}$$

Then, using (6),

$$G(\mathbf{x}) = \frac{1}{n} \cdot \frac{a_1}{r_1}; \quad G(\mathbf{y}) = \frac{1}{n} \cdot \frac{a_2}{r_2}; \quad G(\mathbf{x}') = \frac{1}{n} \cdot \frac{b_1}{t_1}; \quad G(\mathbf{y}') = \frac{1}{n} \cdot \frac{b_2}{t_2};$$
$$G(\mathbf{x}+\mathbf{y}) = \frac{1}{n} \cdot \frac{a_1 + a_2}{r_1 + r_2}; \quad G(\mathbf{x}'+\mathbf{y}') = \frac{1}{n} \cdot \frac{b_1 + b_2}{t_1 + t_2}.$$

Since a_1, r_1, a_2, r_2 and b_1, t_1, b_2, t_2 are such that (5) holds, we have

$$G(\mathbf{x}) > G(\mathbf{x}') \text{ and } G(\mathbf{y}) > G(\mathbf{y}'), \text{ but } G(\mathbf{x}+\mathbf{y}) < G(\mathbf{x}'+\mathbf{y}').$$

In other words, we have an example of the type of paradox given by (1) for the Gini index.

So, given an example of Simpson's paradox, the task boils down to obtaining (non-decreasingly ordered) income distributions $\mathbf{x}$, $\mathbf{y}$, $\mathbf{x}'$ and $\mathbf{y}'$ such that (7) holds. Consider $\mathbf{x}$. The goal is to obtain $x_1, \ldots, x_n$ such that the following two linear equality conditions hold: $\alpha_1 x_1 + \cdots + \alpha_n x_n = a_1$ and $x_1 + \cdots + x_n = r_1$. Additionally, the linear inequality conditions $0 \leq x_1 \leq x_2 \leq \cdots \leq x_n$ have to hold. So, we have a system of linear inequalities. In general, solving a system of linear inequalities is related to the linear programming problem and it is not guaranteed that a solution always exist. As a specific technique, the Fourier-Motzkin elimination technique can be applied to obtain a solution. Such a solution provides the income distribution $\mathbf{x}$. In a similar manner, the other income distributions $\mathbf{y}$, $\mathbf{x}'$ and $\mathbf{y}'$ can be obtained.

Let us now briefly consider the Bonferroni index. From (4),

$$B(\mathbf{x}) \quad = \quad \frac{1}{n\mu_n}\left(\sum_{i=1}^{n}(x_i-\mu_i)\right) = \frac{\beta_1 x_1+\cdots+\beta_n x_n}{x_1+\cdots+x_n} \tag{8}$$

where $\beta_i = 1-H_n+H_{i-1}$ and $H_i = 1+1/2+\cdots+1/i$, is the i-th Harmonic number. Again, for a fixed n, $\beta_1,\ldots,\beta_n$ are constants. Suppose we can find income distributions $\mathbf{x}=(x_1,\ldots,x_n)$, $\mathbf{y}=(y_1,\ldots,y_n)$, $\mathbf{x}'=(x_1',\ldots,y_n')$ and $\mathbf{y}'=(y_1,\ldots,y_n')$ such that

$$\left.\begin{array}{ll} \beta_1 x_1+\cdots+\beta_n x_n = a_1; & x_1+\cdots+x_n = r_1; \\ \beta_1 y_1+\cdots+\beta_n y_n = a_2; & y_1+\cdots+y_n = r_2; \\ \beta_1 x_1'+\cdots+\beta_n x_n' = b_1; & x_1'+\cdots+x_n' = t_1; \\ \beta_1 y_1'+\cdots+\beta_n y_n' = b_2; & y_1'+\cdots+y_n' = t_2. \end{array}\right\} \tag{9}$$

Then

$$B(\mathbf{x}) = \frac{a_1}{r_1};\quad B(\mathbf{y}) = \frac{a_2}{r_2};\quad B(\mathbf{x}') = \frac{b_1}{t_1};\quad B(\mathbf{y}') = \frac{b_2}{t_2};$$
$$B(\mathbf{x}+\mathbf{y}) = \frac{a_1+a_2}{r_1+r_2};\quad B(\mathbf{x}'+\mathbf{y}') = \frac{b_1+b_2}{t_1+t_2}.$$

Since a_1, r_1, a_2, r_2 and b_1, t_1, b_2, t_2 are such that (5) holds, we have

$$B(\mathbf{x}) > B(\mathbf{x}') \text{ and } B(\mathbf{y}) > B(\mathbf{y}'), \text{ but } B(\mathbf{x}+\mathbf{y}) < B(\mathbf{x}'+\mathbf{y}').$$

In other words, we have an example of the type of paradox given by (1) for the Bonferroni index. The task of obtaining the income distributions $\mathbf{x}, \mathbf{y}, \mathbf{x}'$ and $\mathbf{y}'$ satisfying (9) and the non-decreasingly ordered condition is the same as that described for the Gini index.

6 Concluding Remarks

In this paper, we have identified a Simpson-type paradoxical behavior of inequality indices and have demonstrated it for the Gini and Bonfeeroni indices. It remains to be examined whether concrete examples of the paradox can be obtained for all Lorenz consistent inequality indices, in particular, for the generalized entropy family. One way to proceed along this line is to develop a general method to obtain distributions that may lead to the paradox.

References

Aaberge, R.: Gini's nuclear family, Journal of Economic Inequality 5,305–322 (2007).

Bandyopadhyay, P. S., Nelson, D., Greenwood, M., Brittan, G., and Berwald, J.: The logic of Simpson's paradox, Synthese, 181, 185–208 (2011).

Bercena-Martin, E. and Silber, J.: On the generalization and decomposition of the Bonferroni index, Social Choice and Welfare, 41,763–787 (2013).

Blyth, C.: On Simpson's paradox and the sure-thing principle. Journal of the American Statistical Association 67, 364–366 (1972).

Bonferroni, C.: Elemente di Statistica Generale. Libreria Seber, Firenze (1930).

Chakravarty, S.R. and Sarkar, P.: New Perspectives on the Gini and Bonferroni Indices of Inequality, Working Papers 538, ECINEQ, Society for the Study of Economic Inequality, 2020.

Cohen, M. and Nagel, E.: An introduction to logic and the scientific method, New York: Harcourt, Brace and Company, 1934.

Cowell, F.: Inequality and Poverty Measures, in The Oxford Handbook of Well-Being and Public Policy, edited by Matthew D. Adler and Marc Fleurbaey, 2016

Donaldson, D. and Weymark, J.A.: A single-parameter generalization of the Gini indices of inequality, Journal of Economic Theory 22 (1980),67–86.

Grimmett, G. R., and Stirzaker, D. R.:, Probability and Random Processes, 3rd Edition, Oxford University Press, 2001.

Pearson, K., Lee, A., and Bramley-Moore, L.: Genetic (reproductive) selection: inheritance of fertility in man, and of fecundity in thoroughbred race horses, Philosophical Transactions of the Royal Society of London, Series A, 192, 257–330 (1899).

Simpson, E.: The interpretation of interaction in contingency tables, Journal of the Royal Statistical Society, Series B, 13, 238–241 (1951).

Sen, A.K.: On Economic Inequality, Clarendon Press, 1973.

Weymark, J.A.: Generalized Gini inequality indices, Mathematical Social Sciences 1, 409–430 (1981).

Yule, G.: Notes on the theory of association of attributes in statistics, Biometrika, 2, 121–134 (1903).