2025

On a transformation of the Gini coefficient into a well-behaved social welfare function

By ODED STARK

Reprinted from

KYKLOS

Kyklos

WILEY

KYKLOS

ORIGINAL ARTICLE OPEN ACCESS

On a Transformation of the Gini Coefficient into a Well-Behaved Social Welfare Function

Oded Stark[1,2]

[1]University of Bonn, Germany | [2]University of Warsaw, Poland

Correspondence: Oded Stark, ZEF, University of Bonn, Genscherallee 3, D-53113 Bonn, Germany. **E-mail:** ostark@uni-bonn.de

Received: 15 July 2024 | **Revised:** 22 August 2024 | **Accepted:** 30 September 2024

Abstract

Following Sen's (1973) characterization of the Gini coefficient as a ratio between a measure of aggregate income-based stress ("depression" in Sen's terminology) and aggregate income, we transform the Gini coefficient into a social welfare function rather than having the Gini coefficient feature as an input in a social welfare function as in Sen (1973 and 1997), Sen (1976), and Sen (1982). The "Gini social welfare function" assigns weights that reflect preferences to aggregate income and to aggregate income-based stress (income inequality), a desirable property that a social welfare function in which the Gini coefficient features as an input does not have. The transformation bears on the formation of public policy and on the welfare analysis of policy interventions.

Keywords: Aggregate income-based stress; Aggregate income; Gini coefficient; Transformation of the Gini coefficient into a social welfare function

JEL classification: C43; D01; D31; D63; H53; I31; I38; P46

Introducing the Gini coefficient, Sen (1973, p. 33) writes: "In any pair-wise comparison the man with the lower income can be thought to be suffering from some depression on finding his income to be lower. Let this depression be proportional to the difference in income. The sum total of all such depressions in all possible pair-wise comparisons takes us to the Gini coefficient."

1 | Introduction

The starting point of this paper is that the Gini coefficient incorporates two distinct dimensions of income: relative, as per Sen's (1973, p. 33) representation, and absolute. As a measure of a "bad" (given that inequality is deemed a "bad"), the coefficient is "rewarded" (increased) by lower relative income and is "penalized" (decreased) by higher absolute income. In this regard, the coefficient reflects the natural preferences of individuals: when an individual compares his income with the incomes of other individuals and experiences stress ("depression" in Sen's terminology) when the outcome of the comparisons is unfavorable, we can quantify the level of stress by a measure of the sum of the income excesses normalized by the size of the population. Further, when incomes happen to be rather high during the time in which stress is experienced, we can assume that this prevalence of high incomes will assuage the stress being felt. As a ratio of the measure of aggregate stress (denoted below by *AS*) divided by aggregate income (denoted below by *TI*), the Gini coefficient reflects these population preferences.

We draw on the ratio of aggregate stress to aggregate income as a basis. We transform the Gini coefficient to a social welfare function. This conversion is revealing because it goes further than incorporating the Gini coefficient as an input in a social welfare function. Moreover, the "Gini social welfare function" has a desirable property not possessed by a social welfare function in which the Gini coefficient features as an input. This property is the capability to assign weights in the function to aggregate income and income-based stress and adapt these weights to a population's preferences, resulting in a measure of social welfare that duly reflects the manner in which people assess their wellbeing.

2 | Transforming the Gini Coefficient into a Social Welfare Function

To demonstrate the intriguing way in which the characterization of the Gini coefficient as a ratio between a measure of aggregate income-based stress and aggregate income can be drawn upon to yield a social welfare function, we refer first to a population of two individuals with incomes $0 < y_1 < y_2$.

Sen (1973 and 1997), Sen (1976), and Sen (1982) proposed a social welfare function, denoted here by SWF_{SEN}, that he defined as $SWF_{SEN} \equiv \mu(1-G)$, where μ is income per capita, and G is the Gini coefficient. In the case of incomes $0 < y_1 < y_2$, Sen's social welfare function takes the following form:

$$SWF_{SEN} = \frac{y_1+y_2}{2}\left[1-\frac{\frac{1}{2}(y_2-y_1)}{y_1+y_2}\right] = \frac{1}{4}(3y_1+y_2). \tag{1}$$

In the Appendix, we show how the expression of G in (1) as $\frac{\frac{1}{2}(y_2-y_1)}{y_1+y_2}$ is derived as a special case of Sen's (1973) definition of the Gini coefficient. The SWF_{SEN}, Sen asserts, represents an improvement in the measurement of social welfare merely by income per capita. In "defense" of SWF_{SEN}, Sen (1997, p. 137) remarked that "its interpretation as the mean income modified downward by the Gini inequality adds to its attraction as an intuitive and usable welfare indicator." However, the SWF_{SEN} representation is worrisome because *any* increase in the *top* income increases the level of social welfare. Yet the idea of expressing social welfare as a negative function of the Gini coefficient is that an increase in the coefficient should penalize social welfare, and if the increase in the coefficient is large, which happens when a given increase in income happens with the top income, then social welfare should be penalized harshly. This consequence is not adequately yielded by Sen's social welfare function. In Stark and Budzinski (2021), we already raised, but did not resolve, this concern. We were troubled to find that a marginal increase in income of *any* individual, regardless of that individual's position in the income distribution, improves welfare as measured by the $\mu(1-G)$ social welfare function. We hasten to add that we had no issue with an increase in the bottom income, which increases income per capita μ and decreases the Gini coefficient G, thereby increasing social welfare as per Sen's social welfare function.

To correct for the aforementioned unwarranted attribute of Sen's social welfare function, we construct a social welfare function as follows. For the case of incomes $0 < y_1 < y_2$, the population's aggregate or total income, which we denote by TI, is y_1+y_2, and the population's measure of income-based stress, which we denote by AS, is $\frac{1}{2}(y_2-y_1)$. This last term is the income-based stress experienced by individual 1; it is the income gap to which individual 1 is subject, normalized by the size of the population.[1] For $\alpha \in (0,1)$, we define $SWF^* \equiv (1-\alpha)TI - \alpha AS$, which in the present case of incomes $0 < y_1 < y_2$ takes the form

$$SWF^* = (1-\alpha)(y_1+y_2) - \alpha\frac{1}{2}(y_2-y_1). \tag{2}$$

The coefficient α is the weight given to income-based stress, and the coefficient $1-\alpha$ is the complementary weight given to absolute income. In Section 3, we elaborate on the meaning and role of α.

Because

$$\frac{dSWF^*}{dy_1} = 1 - \frac{\alpha}{2} > 0, \tag{3}$$

we obtain that a marginal increase in the bottom income increases social welfare. And because

$$\frac{dSWF^*}{dy_2} = 1 - \frac{3}{2}\alpha, \tag{4}$$

then $\frac{dSWF^*}{dy_2} < 0$ if $\alpha > \frac{2}{3} \equiv \alpha^*$, so we obtain that when the income-based stress that arises from having a low relative income is given a weight that is higher than a critical fraction α^* then, from the perspective of social welfare, an increase in the top income is not beneficial; the resulting increase in the aggregate income-based stress severely penalizes social welfare.

This result is significant. It stands to reason that populations can and do differ in the intensity of their concern over the experience of income-based stress and, correspondingly, in the weight that they will want to assign to a measure of this stress in their social welfare function. It is thus of paramount importance that the social welfare function of a population be endowed with the capability to account for the intensity of this concern. The capacity to accommodate differential weighting is given by SWF^*, but not by SWF_{SEN}.

We next show that manipulation of the Gini coefficient that results in its transformation into a social welfare function yields an outcome that mirrors the outcome obtained from using SWF^*. To this end, we take the following steps. In the case of incomes $0 < y_1 < y_2$, and as already exhibited in (1),

$$G = \frac{\frac{1}{2}(y_2-y_1)}{y_1+y_2}. \tag{5}$$

Thus, $G = \frac{AS}{TI}$.

From (5) it follows that $\frac{1}{G} = \frac{y_1 + y_2}{\frac{1}{2}(y_2 - y_1)}$, so that $\log\frac{1}{G} = \log(y_1 + y_2) - \log\frac{1}{2}(y_2 - y_1)$. Defining $G_\alpha \equiv \frac{AS^\alpha}{TI^{1-\alpha}}$ for $AS^\alpha = \left(\frac{1}{2}(y_2 - y_1)\right)^\alpha$ and $TI^{1-\alpha} = (y_1 + y_2)^{1-\alpha}$, we obtain, for a population of two individuals,

$$\log\frac{1}{G_\alpha} \equiv (1-\alpha)\log(y_1+y_2) - \alpha\log\frac{1}{2}(y_2-y_1)$$
$$= (1-\alpha)\log(y_1+y_2) - \alpha\log\frac{1}{2} - \alpha\log(y_2-y_1)$$

and

$$SWF^{**} \equiv \log\frac{1}{G_\alpha} = (1-\alpha)\log(y_1+y_2) - \alpha\log\frac{1}{2} - \alpha\log(y_2-y_1),$$

where $\alpha \in (0, 1)$ is as introduced earlier.

We obtain two results:

$$\frac{dSWF^{**}}{dy_1} = \frac{1-\alpha}{y_1+y_2} + \frac{\alpha}{y_2-y_1} > 0 \tag{6}$$

and

$$\frac{dSWF^{**}}{dy_2} = \frac{1-\alpha}{y_1+y_2} - \frac{\alpha}{y_2-y_1} = \frac{y_2-y_1-2\alpha y_2}{(y_1+y_2)(y_2-y_1)}. \tag{7}$$

In (7), because the denominator of the rightmost fraction is positive, $\frac{dSWF^{**}}{dy_2}$ is negative if $y_2 - y_1 < 2\alpha y_2$, that is, if $\alpha > \frac{y_2 - y_1}{2y_2} = \frac{1 - \frac{y_1}{y_2}}{2} \equiv \alpha^{**}$, which is a feasible condition because α^{**} is a fraction. Similarly to what we obtained from the derivative in (4), here too we see that when the income-based stress that arises from having a low relative income is given a weight that is higher than a critical fraction, which in the current case is α^{**}, then, from the perspective of social welfare, an increase in the top income is not beneficial; the resulting increase in the aggregate income-based stress severely penalizes social welfare.

Thus, upon use of SWF^{**}, we obtain results that mirror the results obtained upon use of SWF^{*}: (6) is equivalent to (3), and (7) along with $\alpha > \alpha^{**}$ is equivalent to (4) along with $\alpha > \alpha^{*}$.

For the general case of n individuals, we define (for the sake of clarity of distinction, namely to underscore the fact that we are departing from the case of a population of two individuals, we next add to G_α, AS, TI, and SWF^{**} a superscript n):

$$G_\alpha^n \equiv \frac{(AS^n)^\alpha}{(TI^n)^{1-\alpha}}$$

for $AS^n = \frac{1}{n}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}(y_j - y_i)$, and $TI^n = \sum_{i=1}^{n} y_i$. We now define a social welfare function

$$SWF^{**n} \equiv \log\frac{1}{G_\alpha^n} = (1-\alpha)\log TI^n - \alpha\log AS^n.$$

From our preceding considerations we know that a marginal increase in the income of the "poorest" individual leads to a decrease in AS and an increase in TI and, thus, to a decrease in G_α^n and, therefore, to an increase in SWF^{**n}. Also, because

$$\frac{dAS^n}{dy_n} = \frac{n-1}{n}, \text{ and } \frac{dTI^n}{dy_n} = 1,$$

then

$$\frac{dSWF^{**n}}{dy_n} = \frac{1-\alpha}{TI^n} - \frac{\alpha(n-1)}{nAS^n},$$

so $\frac{dSWF^{**n}}{dy_n}$ is negative if $(1-\alpha)nAS^n < \alpha(n-1)TI^n$, which can be transformed into $\alpha > \frac{nAS^n}{(n-1)TI^n + nAS^n} \equiv \alpha^{**}(TI^n, AS^n)$. This is a feasible condition because $\alpha^{**}(TI^n, AS^n)$ is a fraction. Thus, once again, we obtain that when the income-based stress that arises from having a low relative income is given a weight that is higher than a critical fraction, here $\alpha^{**}(TI^n, AS^n)$, then, from the perspective of social welfare, an increase in the top income is not beneficial.

For $n = 2$, $\alpha^{**}(TI^2, AS^2) = \frac{y_2 - y_1}{2y_2} = \frac{1 - \frac{y_1}{y_2}}{2} = \alpha^{**}$, as before.

3 | Discussion and Conclusions

The insight gained in this paper from an exploration of the inner workings of the Gini coefficient as a composite measure of aggregate income-based stress and aggregate income is relevant for the formation of public policy and for the welfare analysis of policy interventions. Given that the core business of a government is to raise welfare, and given that people's wellbeing is determined both by relative income and absolute income, it is of paramount importance that a government has in place a social welfare function that meets these people's preferences. We showed that, in and of itself, the Gini coefficient can be transformed into a social welfare function that meets this criterion, and that this function is free from a drawback of a social welfare function that is constructed using the Gini coefficient as an input. To the best of our knowledge, this intriguing conversion was not identified previously.

People can differ in their tolerance for income-based stress arising from income inequality as a result of cultural and historical factors. For example, people in communist societies or people who grew up and lived for a long time in communist societies will be less used to and be more annoyed by large income inequalities and so exhibit a high α coefficient.

The incorporation in SWF^{*} of the coefficient α enables a government to fine-tune policies in response to shifting preferences of its population. For example, it could be argued that when incomes are generally low, people become quite upset at also having low relative incomes (the coefficient

α is relatively high), whereas when incomes are generally high, people are more tolerant about experiencing low relative incomes (the coefficient α is relatively low). Alternatively, it could be argued that when incomes are generally low, people care more about absolute income and less about relative income (the coefficient $1-\alpha$ is relatively high, the coefficient α is relatively low) than when incomes are generally high (the coefficient $1-\alpha$ is relatively low, the coefficient α is relatively high).

To reiterate, the condition $\alpha > \alpha^{**}$ states that when the income-based stress that arises from having a low relative income is given a weight that is higher than a critical fraction α^{**}, then, from the perspective of social welfare, an increase in the top income is not beneficial; the resulting increase in the aggregate income-based stress severely penalizes social welfare. Consider, then, $\alpha^{**} = \frac{y_2 - y_1}{2y_2}$. Because α^{**} is increasing in y_2, it is more difficult for the condition $\alpha > \alpha^{**}$ to hold when y_2 is higher. Therefore, in a richer society, meaning that the top income y_2 is higher, there is a broader scope for the income of the richest individual to increase without SWF^{**} taking a beating on account of experiencing aggregate income-based stress.

Calculating SWF_{SEN}, for example as in (1), requires obtaining income data y_1 and y_2, whereas calculating SWF^{*}, for example as in (2), requires obtaining both income data y_1 and y_2 and taste datum α. This additional requirement need not be insurmountable. A survey can yield an estimate of the rate of substitution between income-based stress and absolute income. (Upon assuming similarity / homogeneity in tastes, the individuals in a given population can be taken to share the same α.) That is, given y_1 and y_2, we can ask, when y_2 increases, by how much y_1 should increase so as to retain the same level of welfare that existed prior to the increase in y_2. This will yield an equation with a single unknown α, thereby enabling the calculation of SWF^{*}.

In 1912 Corrado Gini constructed an index of dispersion without any link or reference to the measurement of inequality.[2] That index was subsequently transformed by economists and others into a widely used measure of inequality.[3] Perhaps we could refer to that usage as the first Gini index conversion. It remains to be seen whether economists and others will be inclined to refer to the transformation reported in this paper as the second Gini index conversion.

Appendix

In population $N = \{1, 2, \ldots, n\}$, $n \geq 2$, let $y = (y_1, \ldots, y_n)$ be the vector of incomes of the members of the population, and let these incomes be ordered as follows: $0 < y_1 < y_2 < \ldots < y_n$. As per Sen (1973), the Gini coefficient can be presented as

$$G \equiv \frac{\sum_{j=1}^{n}\sum_{i=1}^{n}\left|y_i - y_j\right|}{2n^2\bar{y}}, \tag{8}$$

where $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$ is the population's average income.

On noting that $\sum_{j=1}^{n}\sum_{i=1}^{n}\left|y_i - y_j\right| = 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}(y_j - y_i)$, an equivalent representation of G in (8), which dispenses with the need to operate with absolute values, is

$$G = \frac{\frac{1}{n}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}(y_j - y_i)}{\sum_{i=1}^{n} y_i}. \tag{9}$$

When $n = 2$, (9) reduces to (5).

As an aside, we note that the right-hand side of (9) is an explicit expression of $\frac{AS^n}{TI^n}$ used in Section 2.

Acknowledgments

I am indebted to a reviewer for illuminating comments and kind words and to David Stadelmann for advice and guidance. Open access funding enabled and organized by Projekt DEAL.

Endnotes

[1] AS stands for aggregate stress (the sum of the levels of stress, that is, the sum of the income-based gaps of individuals' incomes). In the case of a population of two individuals whose incomes are distinct, there is only one income-based gap, so use of the term aggregate is not needed. Use of the term aggregate is appropriate when the population consists of three or more individuals.

[2] Ceriani and Verme (2012) present an illuminating account of the thinking that led Gini (1912) to formulate his index.

[3] "[T]he Gini coefficient [is] still the most commonly used measure of inequality in empirical work." (Sen 1973, p. 149).

References

Ceriani, Lidia and Verme, Paolo (2012). "The origins of the Gini index: Extracts from Variabilità e Mutabilità (1912) by Corrado Gini." *Journal of Economic Inequality* 10: 421–443.

Gini, Corrado (1912). *Variabilità e mutabilità. Contributo allo Studio delle Distribuzioni e delle Relazioni Statistiche*. Bologna: C. Cuppini.

Sen, Amartya (1973 and 1997). *On Economic Inequality*. Oxford: Clarendon Press.

Sen, Amartya (1976). "Real national income." *Review of Economic Studies* 43(1): 19–39.

Sen, Amartya (1982). *Choice, Welfare and Measurement*. Oxford: Blackwell.

Stark, Oded and Budzinski, Wiktor (2021). "A social-psychological reconstruction of Amartya Sen's measures of inequality and social welfare." *Kyklos* 74: 552–566.