On an Apparently Innocuous Difference in Two Versions of Gini's Coefficient of Inequality: A Symposium[†]

Rajiv Sethi
Debraj Ray
Samuel Bowles and Wendy Carlin

October 2022

1. Introduction

The Gini coefficient is a leading measure of inequality, and is widely employed in the social sciences as a summary statistic for income or wealth distributions. This article combines three essays by the authors—a conversation, really—about the foundations of the Gini coefficient. Our differences, expressed within the boundaries of the congenial esteem that we have for one another, center around an appropriate formula for the Gini.

To the immediate query as to why we simply don't use Gini's own formula and call it a day, there are two responses. First Gini himself proposed no less than thirteen versions of his measure, so that avenue isn't of much help. Less flippantly, our discussion is not so much about the choice of a specific formula. Rather, it is about what we are trying to capture when me measure inequality, and the appropriate role of a particular philosophical axiom that is often involved in inequality measurement—the population principle.

We should note that our conversation is already anticipated, in part, by scholars who have thought about these and related issues far more deeply than we have; among them, John Creedy, Anthony Shorrocks and S. Subramanian (see, for instance, Subramanian [2010]).

2. Notes on the Gini Coefficient, by Rajiv

The Gini coefficient has been introduced to generations of students using some variant of the formula:

$$G = \frac{1}{2\mu n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} |y_i - y_j|,$$

where $y_1, ..., y_n$ are non-negative income levels in a population of size n, and $\mu > 0$ is mean income in this population [Sen, 1973, Ray, 1998]. This index has an appealing interpretation as the average absolute difference between all pairs of individuals, relative to the mean income in the population.

[†]Sethi: Barnard College, Columbia University and the Santa Fe Institute, rs328@columbia.edu; Ray: New York University and the University of Warwick, debraj.ray@nyu.edu; Bowles: Santa Fe Institute, samuel.bowles@gmail.com; Carlin: University College London, w.carlin@ucl.ac.uk. Thanks to Julia Schwenkenberg and S. Subramanian for useful comments. The original conversation is archived at https://debrajray.com/2021/10/01/a-discussion-of-the-gini-coefficient/.

But the pairs in this case include individuals paired with themselves, with the corresponding income differences being identically zero. The numerator is unaffected by their inclusion, but the denominator is inflated relative to the case when such pairs are excluded.

This reasoning has led some to favor an alternative version of the index that simply excludes self-matched pairs, and hence involves just $n(n-1)$ rather than n^2 comparisons [Jasso, 1979, Deaton, 1997, Bowles and Carlin, 2020]. The resulting measure of inequality is

$$G' = \frac{1}{2\mu n(n-1)} \sum_{i=1}^{n} \sum_{j=1}^{n} |y_i - y_j| \, .$$

Clearly

$$G = \frac{n-1}{n} G'$$

so the two measures are close for large n and converge in the limit. But for small populations the difference between the two can be substantial.

Both G and G' have been axiomatized by Thon [1982]. Four principles together yield G, namely the transfer principle (all order-preserving and equalizing transfers reduce inequality), population symmetry (pooling identical populations results in the same level of inequality as in the component populations), constant population comparability (at any given population, the range spanned by the index should not depend on total income), and equidistance (all order-preserving and equalizing transfers between people at adjacent income levels have the same effect on the index).

To obtain an axiomatization for G', one can dispense with population symmetry and strengthen the comparability axiom to strong comparability, which requires that the range spanned by the index depends neither on total income nor on population. As Thon [1982, p. 140] puts it: ”One might indeed want to postulate that the range of an inequality index is to be the same over the redistribution of any total income not only between a given number of people but between any number of people.”

So one way to decide on whether G or G' is preferred is to consider which of the two axioms—population symmetry or strong comparability—one is more willing to discard. Allison [1979] makes a case for G on the grounds that reasonable measures ought to satisfy population symmetry. That is, pooling two (or more) identical populations should result in the same level of inequality in the pooled population as existed in the component groups. If one person in a group of two has all the income, measured inequality should be the same as if all income was shared equally by a thousand people in a group of two thousand. This is the case with G (which equals one-half in each case) but not with G' (which equals one in the former case and is close to one-half in the latter).

One could argue, however, that the pooling of two or more identical populations could well result in a composite that is qualitatively different, and can reasonably be held to have a very different level of inequality. In the example above, there is a group of people in the pooled population who share equality, who must accommodate each other in social and political life, and who may establish rights and responsibilities that apply only to themselves but nevertheless operate as a constraint on behavior and may eventually spread to society more broadly. Some will consider these arguments extraneous and irrelevant, but they may be persuasive to others.

By the same token, one can construct examples that seem to suggest that G' is a poor measure of inequality. It assigns the same value to a group in which one of two people has all income as it does to a group in which one of a two thousand does ($G' = 1$ in each case, while G is one-half in the former and close to one in the latter). Worse, it assigns greater inequality to the former society (in which only one of two people has positive income) than it does to a society in which just two people out of a million have positive incomes. It could be argued that income is far more concentrated when the elite is small relative to the total population, and that this ought to be reflected in the inequality measure.

There is another important argument to be considered, which is the consistency of the measures with the partial order on arbitrary income distributions induced by the Lorenz criterion.

Any income distribution associated with a finite population can be represented by a set of points in two dimensions, with the cumulative share of the population on the horizontal (in order of increasing income) and the cumulative share of income on the vertical. A Lorenz curve is obtained by interpolating these points to get a non-decreasing and convex function on the unit interval.[1] If all incomes are equal, there is only one interpolation that satisfies these criteria—the line of perfect equality or identity function. All other distributions yield curves that lie below this line, meeting it at the end points.

The process of interpolation associates with each finite population distribution an infinite population counterpart. This allows us to compare distributions regardless of total income or population, by simply comparing their Lorenz curves—if a curve lies closer to the line of equality at all points it corresponds to a distribution with lower inequality.[2] The result is a partial order on the set of all income distributions.

However, the particular partial order thus obtained clearly depends on the method of interpolation. Given two distributions, one method of interpolation may provide a clear ranking, while a second may involve intersecting curves. So when one speaks of a Lorenz order, or consistency with the Lorenz criterion, there has to be a method of interpolation either explicitly or implicitly assumed.

The assumed interpolation is usually piecewise linear [Ray, 1998]. This always generates a curve that has the properties necessary for interpretation as an income distribution in an infinite population. In addition, it is the only method of interpolation that respects the population symmetry axiom. That is, with piecewise linear interpolation, the merging of two identical populations results in a distribution which lies on the (interpolated) Lorenz curve corresponding to the component populations.

[1] Lorenz [1905] placed population shares on the vertical axis and income shares on the horizontal, resulting in concave curves.

[2] When comparing two distributions with the same population size, clearly no interpolation is required.

The standard Gini coefficient G is a completion of the partial order generated by piecewise linear interpolation. Specifically, G is the ratio of the area between the line of perfect equality and the Lorenz curve thus constructed, and the total area below the line of perfect equality. It is in this sense that G is Lorenz consistent.

However, there exist several methods for nonlinear interpolation that can generate Lorenz curves with all the required properties [Gastwirth and Glauberman, 1976, Cowell and Mehta, 1982]. These methods have been developed to deal with empirical applications involving binned data, but can also be applied to the case when we have data at the individual level for a finite population. A key step involves the fitting of an underlying density function to the available data points. The piecewise linear interpolation corresponds to a piecewise uniform density. Other densities map on to other Lorenz curves, including curves that are constructed to be continuously differentiable.

In his response to this note, Debraj Ray makes the important point that G' is inconsistent with the standard Lorenz ranking (based on piecewise linear interpolation). One can go further—there is no method of interpolation consistent with convexity and the other required properties that generates a partial order with which G' is consistent. To see this, consider any method of interpolation with the necessary properties. Corresponding to this, there will be some continuous Lorenz curve associated with the two person distribution in which one person has all income. By choosing a population n to be sufficiently large, and considering a distribution in which just two people in this population share all income equally, one can get a Lorenz curve that lies strictly below the first one for the same method of interpolation. This will be treated as more unequal under the Lorenz criterion (based on the chosen interpolation). But under G' the former has maximal inequality while the latter does not.

Although G' is inconsistent with the Lorenz criterion (for any interpolation), it does have an interesting geometric interpretation. Suppose that one uses a step function for interpolation rather than a continuous convex function. In this case the ”line of perfect equality” is replaced by a step function that lies strictly below the conventional perfect equality line for finite populations, and is sensitive to the size of the population, approaching the conventional line in the limit. In addition, there is a line of perfect *inequality* that lies on the horizontal axis, and is independent of population size.

Unlike the conventional Lorenz curve, this step function interpolation (being non-convex) cannot be interpreted as an income distribution in an infinite population. Nevertheless, one may ask whether the area between the step function corresponding to perfect equality and the step function corresponding to the observed income distribution can be used as a measure of inequality that ranks all distributions regardless of total income or population size. Indeed it can, and the ratio of this area to the total area below the perfect equality step function is precisely equal to G'. To see this, one need only verify that this area measure satisfies strong comparability, since it clearly satisfies equidistance and the transfer principle. And this is also clearly true—the ratio must be zero when income is equally distributed, and must be one when a single individual has all income.

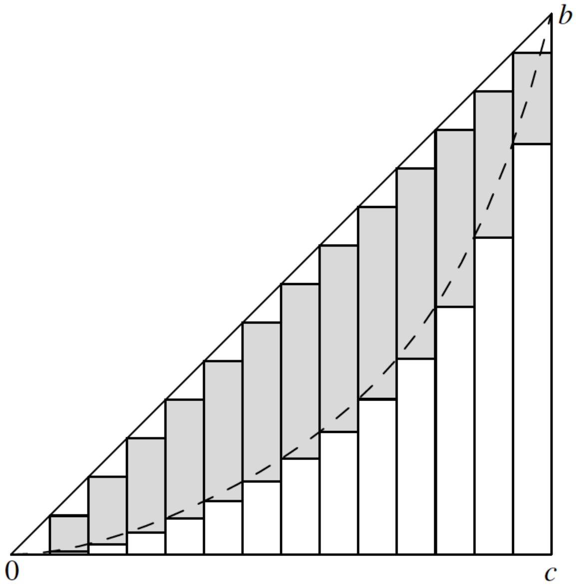

Figure 1. Geometric interpretation of G' in Gini [1914].

In fact, the equivalence of G' and this particular ratio of areas was recognized by Gini himself, and is self-evident from Figure 1, which appears in Gini [1914].[3] This figure depicts G' exactly, based on Gini's own definition, for a population with $n = 14$ and a particular distribution of income. Notice that Gini uses a smooth, nonlinear interpolation (the dashed line) to construct a ”Lorenz curve” for this finite population. This curve fails to satisfy the population symmetry axiom—merging two identical fourteen-person populations of this kind would result in points that do not lie on the curve as drawn.[4]

Finally, consider which of the two measures has a stronger claim to be known as the Gini coefficient. As it happens, both measures may be found in Gini [1912], although he appears to favor G' as a measure of income inequality [Ceriani and Verme, 2012]. And in Gini [1914], translated and published as Gini [2005], he is explicit about the requirement of strong comparability, stating that for any population size, his index of concentration ”ranges from 1, in the case of perfect concentration, to 0, in the case of equidistribution.”

Nevertheless, as Allison [1979] has observed, ”both versions of the Gini index have found their way into the statistical literature, and neither one can be said to be incorrect.” It is probably best if students are at least made aware of the existence and historical origins of both, and presented with the arguments in favor of each. There is no uniquely correct Gini coefficient.

[3] The figure shown here is taken from the translation [Gini, 2005]; I thank Sam Bowles for bringing it to my attention.
[4] in order for it to satisfy population symmetry it would have to be piecewise linear, as noted above.

3. NOTES ON "NOTES ON THE GINI COEFFICIENT," BY DEBRAJ

A point of historical interest—undoubtedly arcane to some—is that the worthy Corrado Gini (Gini 1914) produced no fewer than 13 Gini coefficients for our perusal. I'm not entirely sure what all thirteen are but two of them seem to have survived the test of time and are widely used. The first is the formula

$$G = \frac{1}{2\mu n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} |y_i - y_j|,$$

where the n stands for population, the y's are incomes, and μ is their mean. The second is given by the formula

$$G' = \frac{1}{2\mu n(n-1)} \sum_{i=1}^{n} \sum_{j=1}^{n} |y_i - y_j|,$$

a seemingly minor difference, dividing as it does by $n(n-1)$ instead of n^2. So G and G' agree ordinally on all comparisons *within* any given population, while they could disagree across comparisons with varying populations.

Is this worth a sleepless night or two? Not really, especially if you're not devoted to teaching the stuff with the care and precision which my erstwhile student (and now friend) Julia Schwenkenberg brings to it. Julia was teaching her Rutgers students the fundamentals of inequality measurement and came across the following example in the CORE Open Access textbook:[5]

> "There are two people in the society and one has all the income . . . [This is] perfect inequality, as you would expect,"

going on to observe that the Gini (or perhaps I should say, *the* Gini in their view) was therefore equal to its maximum value of 1: a foregone conclusion, seemingly.[6]

But was it? Julia was well aware of the "population neutrality" principle underlying inequality measurement, which stated as an axiom that population cloning of all individuals — while keeping their incomes unchanged — "should not" change any Lorenz-consistent measure of economic inequality. So the configuration in the text should exhibit no change in inequality if, instead, two people had one unit each of income and the other two had none. But the latter configuration is surely *less* unequal than one in which one person had two units of income and the other three had none. Ergo, the very first situation should be less unequal than one in which one person out of four (as opposed to one of out of two) had all the income.

And yet: if we agree with that, should there not be room for the Gini coefficient to rise some more from its two-person value of 1? What is it doing at 1 already?

[5]For more detail on the CoreEcon project, see https://www.core-econ.org/.
[6]See https://bit.ly/3kB8KNN.

This was intriguing enough (and confusing enough for all concerned) that Julia wrote to me about it, especially since my textbook (Ray 1998) and landmark monographs such as Sen [1973] use the formula for G, whereas the CORE text seemed to be using another formula altogether. After discussions with Rajiv Sethi, and later Sam Bowles and Wendy Carlin, it soon emerged that "the" Gini in the CORE text was G', that it was a version favored by two of the CORE authors (Bowles and Carlin 2020), and indeed, that it hits its maximum value in the example above. The other one, G, does not: it equals 1/2, well below its maximum value of 1.

Rajiv Sethi's excellent notes on the subject (Section 2), to which this is a response, lay bare the difference. As already noted, both measures (and eleven others to boot) had been proposed by the prolific Gini, so no claim to true inheritance could be advanced on that somewhat legalistic but otherwise useless basis, "what did Gini really say?" — fortunately so, for many truths have been trampled underfoot by such convenient excuses. Rather, we need to truly evaluate the measures from first principles to advance the discussion, which is what Rajiv does in large part.

The measure G satisfies all the axioms that underly the Lorenz partial order; namely, population and income invariance, as well as the transfers principle. The Lorenz order, in turn, is at the very heart of inequality measurement and forms a welfare basis for it — see, among others, the work of Atkinson [1983] and Dasgupta et al. [1973].[7] G completes this partial order — see Thon [1982] for an axiomatization. It is, to be sure, not the *only* complete order that completes the Lorenz: other examples include the coefficient of variation or the Theil index — see Ray [1998] for more on these matters. But G is one of them; it is *Lorenz-consistent* in the sense of satisfying the axioms that I've just mentioned.

On the other hand, as Rajiv explains, G' would like to highlight the fact that in the two-person example given above, inequality has been stretched to its limit: how can things be any more unequal *in that two-person society*? This brings us to what Rajiv, following Thon [1982], calls strong comparability: "the range spanned by the index depends neither on total income nor on population." In particular, within any society with a fixed population, the index should be able to move from its minimum value (0) to its maximum (1), so that the index recognizes clearly the upper limits to inequality *for that fixed population size*. Clearly, G does not do this — it varies only from 0 to 1/2 in a two-person society. Indeed, by the argument given five paragraphs ago, we can already see that *no measure satisfying the population principle and the transfers principle can be strongly comparable*. In a gesture of inclusiveness, Rajiv concludes:

> "It is probably best if students are at least made aware of the existence and historical origins of both, and presented with the arguments in favor of each. There is no uniquely correct Gini coefficient."

[7] The characterization by Rothschild and Stiglitz [1970] of "increasing risk" is also relevant here, despite its focus on risk and uncertainty; the two have parallel features.

Given a choice between two incompatible desiderata, I can sympathize with Rajiv's assertion that there is no "uniquely correct" Gini. Indeed, given the enormous number of Lorenz completions at our disposal, *there is no uniquely correct measure of inequality even under the Lorenz axioms*, let alone a uniquely correct Gini. Every measure must be evaluated by the core ethical axioms they satisfy, and we then need to reach into our own ethical system to see which set of axioms fit the best. So in this sense, I agree with Rajiv.

That said, there are axioms and there are axioms. I have already mentioned the long and venerable history of the Lorenz curve, which goes back to Lorenz [1905]. Its foundation is laid by a fundamental set of axioms: population neutrality, income neutrality, as well as the transfers principle of Pigou and Dalton. These axioms can and have been questioned; for instance, my work with Joan Esteban on the measurement of polarization (Esteban and Ray 1994a) comes from dropping the transfers principle. But as contributions to a *welfare economics* foundation for inequality measurement, these are the key axioms, and all further explorations begin from them — or should.

The fact that G satisfies all the three axioms, while G' as already noted fails population neutrality, is a priori (though not yet definitive) cause for suspecting the credentials of G'. A noteworthy example comes from Foster [1983], who writes down all the axioms (or "properties") that underpin Lorenz *except for population neutrality*, and then observes of the rest:

> "In fact, since each property is a restriction . . . in isolation without reference to crosspopulation comparisons, [they] admit even more measures than Fields and Fei [1978] indicate. Consider the measure that takes the Gini coefficient index at even sized populations, and the coefficient of variation index at odd. *This absurd measure* quite clearly satisfies all [the] properties.
>
> The above example serves to point out the desirability of a property which would coordinate the indices *into one cohesive measure*. Another property [the population principle] suggested by Dalton does this in a particularly natural way" (emphases mine).

As Foster correctly notes, the population principle is imposed to make the reader aware that we cannot have potentially unrelated and therefore "absurd" measures on different population layers: they need to be connected in a "cohesive" way. Foster goes on to introduce the very population neutrality principle — due to Dalton and embedded in the Lorenz curve — that G satisfies and G' does not, the latter because it refuses to entertain the requirement that its comparisons within a population must be contextualized in the space of all populations.

To see directly the failure in the cohesion of G', consider again the CORE example, in which one person gets all the income in a two-person society. According to G' (and CORE), this situation *is just as unequal* as one in which one person gets all in the income in a million-person society. To me, this is indeed an example of the absurdity that Foster refers to when different population layers are not connected in any coherent way. Taking the example a step further, we must conclude that under G', a situation in which two persons share all the income in a million-person society is *strictly more equal* than the two-person example in the text. This even more absurd consequence comes from the additional application of the transfers principle, which G' does satisfy.

Of course, G exhibits none of these strange behaviors. It would rank the one-in-two example as more equal than the one-in-a-million example, and ditto with the two-in-a-million example.

Rajiv would respond that G fails strong comparability: it does not hit 1 when the two-person society is stretched to its unequal limit. Well, I don't see why strong comparability makes sense. Why must a measure declare perfect inequality just because feasibility constrains a particular situation from exhibiting still greater inequality? The latter is a property of the feasible set and should not influence evaluation, just as a utility function is not affected by the budget set on which it operates. To explain why G' fails this desideratum, consider a two-person society in which either person can asexually clone itself into two, with further income transfers possible among or across clones. Such cloning is prohibited, and so all we can do in this society is transfer income across two individuals without cloning either of them. Strong comparability then requires an inequality measure to hit its maximum possible value when one person has all the income. But now suppose that a divine decree permits cloning. Then the range of inequality values in the earlier two-person range would need to artificially contract so as to accommodate the newly unequal possibilities that arise. In short, this is a case of our measurement indicator responding to the feasible set.

In all of this, I am aware that my arguments do not constitute a logical attack on G' in favor of G, in the sense of claiming that there is some failure of Aristotelian logic in the very fabric of G'. There is no such attack. But it is an appeal to the reader's intuitive sensibilities via the discussion of axioms. In fact, without an axiomatic system, anything goes: it's a veritable free-for-all. We would be taking all too literally Sen's beautiful dedication to his daughters, at the start of *On Economic Inequality*:

> "In the hope that they will find less of it, no matter how they choose to measure it."

In a free-for-all, that utopian dream cannot happen, but once constrained by the spirit of a reasonable axiomatic system, it can. That is the spirit in which I reject G' in favor of G.

4. The Gini Coefficient as a Measure of Experienced Inequality, by Sam and Wendy

Thanks to Rajiv Sethi and Debraj Ray for their stimulating and clarifying memos. Our paper [Bowles and Carlin, 2020] and the notes below present an alternative way of looking at the Gini coefficient(s). Before turning to the substantive issues we begin with three pieces of background about how we came to work on this, about Gini and his coefficient, and about ethics and the measurement of inequality.

4.1. **Background.** Our colleague in the CORE project, Antonio Cabrales, gave his students in the introduction to economics course at University College London the conventional definition of the Gini coefficient, namely (here and throughout using the notation of our 2020 paper for consistency),

$$G^L = \frac{\sum_{i=1}^{i=n} \sum_{j=1}^{j=n} |y_i - y_j|}{2n^2 \underline{y}} \tag{1}$$

along with the usual summary (following Gini) that its value is 1 if one person has all the wealth and zero if wealth is equally distributed. The students could not solve the toy example problems with small n that Antonio had assigned them. After a certain amount of 'check your calculations' and a lot of head scratching we realized that of course they could not make the problems work because G^L is not 1 when one person (in a finite population) has all the wealth; for example, in the case where $n = 2$, $G^L = 0.5$.

Our head scratching included a return to Gini's original paper [Gini, 1914]. There he defined what he called his "concentration ratio" as the sum of the absolute differences among the (unique non-identical) pairs, which we call Δ, divided by the total number of such pairs, relative to mean wealth, $\underline{y}$, multiplied by one half. Hence,

$$\Delta \equiv \sum_{i=j+1}^{i=n} \sum_{j=1}^{j=n-1} |y_i - y_j| \text{ so} \tag{2}$$

$$G = \frac{\Delta}{n(n-1)/2} \frac{1}{\underline{y}} \frac{1}{2} = \frac{\Delta}{n(n-1)} \frac{1}{\underline{y}} \tag{3}$$

from which we can see that G is the mean difference among all pairs in the population (the first term in the expression in the middle of equation 1.2) divided by the mean value of y, giving us the "relative mean difference," times one half. A feature of this measure is, Gini pointed out, that it satisfies the condition that it varied from one (that he termed "maximum concentration") to zero ("minimum concentration") as also shown by [Deaton, 1997]. This representation of the Gini coefficient is consistent with the clever expression for the share of the cake going to the least well off person in the cake-cutting game of Subramanian [2002], of which we were unaware when we wrote our 2020 paper.

Gini showed that this quantity is equal to the area between the (then newly invented) Lorenz curve and the perfect equality line divided by one-half in an infinite population. He provided the appropriate step functions for both the perfect equality "line" and the Lorenz curve for an example of a finite $n = 14$ population. Consistent with our paper, we designate the Lorenz curve-based spatial version of the Gini coefficient as G^L and Gini's version of the concentration ratio in his equation 11 (from his 1914 paper, which is also our equation 1.3 above) as G. As Debraj pointed out, Gini also provided alternative measures (and many more have been proposed since). In their notes, Debraj and Rajiv use G' to refer to our G, and G (without superscript), to refer to our G^L .

Maybe the differences that we have aired in these memos stem from our differing perspectives on the relationship between one's ethics and the choice of an appropriate measure of inequality. Debraj writes "Every measure must be evaluated by the core ethical axioms they satisfy, and we then need to reach into our own ethical system to see which set of axioms fit the best."

Referring to the Lorenz curve Debraj writes: "Its foundation is laid by a fundamental set of axioms: population neutrality, income neutrality, as well as the transfers principle of Pigou and Dalton... as contributions to a welfare economics foundation for inequality measurement, these are the key axioms, and all further explorations begin from them—or should."

This is one way to choose among competing measures. But we have been motivated, instead, primarily by the desire to measure inequality as it is experienced by the members of a society, hence our term for G: experienced inequality. This could be the basis of a normative evaluation; surely one's ethical stance on some distribution of wealth cannot be indifferent to how it is experienced by members of the society. But it might also be an entirely descriptive measure for understanding such things as subjective wellbeing, stress, political attitudes, and the like.

Debraj writes "there are axioms and there are axioms." Our view of experienced equality suggests one to add to the list. **Axiom: Economic inequality is social, that is, it is a relationship between or among people.**

5. Experienced Inequality

We all agree that there is no single right way to measure inequality. Which measure one uses depends of course on the question for which the inequality measure is to provide an answer. This often turns on what it is about inequality that one wishes to capture. In our paper we cite the polarization index due to Joan Esteban and Debraj as an example of a different measure of inequality developed for a specific purpose, to illuminate social conflict [Esteban and Ray, 1994b]. We also cited one of Sam's coauthored works on the polygyny threshold (from anthropology), which shows that the Gini coefficient (either variant) fails to capture the expected relationship between wealth inequality and polygyny, suggesting the need for an alternative measure.

In our paper, we illustrated the Gini coefficient, that is, G, as a statistic describing inequality on a complete network, the edges of which are the differences in wealth between all pairs of the network nodes. We found it helpful both in teaching and in testing our own intuitions to let inequality be about the edges of the network rather than the nodes, that is, about pairwise differences in wealth, not how much wealth each individual has. An example of the two approaches is in Figure 1, with the inequalities counted in equation 1.3 on the left and equation 1.1 on the right.

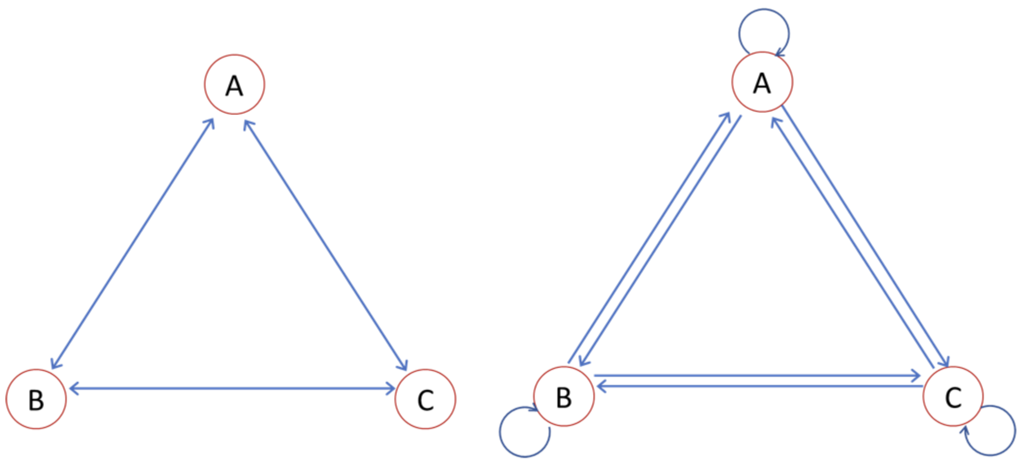

Figure 2. Experienced differences (left panel) as the basis for G and the edges used in the conventional Lorenz curve-based measure G^L (right panel).

If the nodes A, B, and C in Figure 1 have wealth 10, 4, and 3, the Gini coefficient given by equation 1.3 using the network representation in the left panel is 0.412. Using the network representation on the right (that is, equation 1.1), however, the Gini is estimated as 0.274, which needs to be multiplied by $n/(n-1)$ or 1.5 to get the Gini coefficient restricted to the differences among actual pairs in the population (excluding the three "self-on-self" zero differences), as pointed out by Yitzhaki and Schechtman [2013].

We made the case that represented in this way, G captures important aspects of how inequality is experienced by the members of a society in which people are aware of the wealth levels of everyone else. We also find the approach insightful as it allows us to measure experienced inequality in a society in which the relevant comparison set of individuals is not all others, but instead all others to whom one is connected in the network. For example, in a star network with a wealthy individual at the center, there is much more experienced inequality than in a complete network for the same set of nodes and levels of wealth at each node.

To sharpen our intuitions about inequality seen as pairwise differences among members of a population, suppose that every day, individuals are randomly paired to interact—economically, socially, in religious observance, and so on—with another member of the society. In the complete network representation, one of each individual's edges is selected at random. We are interested in the frequency over a great many such random pairings with which a member of the population interacts with a person of similar or different wealth, as we believe that the nature of these interactions will differ in important ways if the wealth differences are significant.

To explore looking at inequality this way, let's consider an economy with just two wealth levels which, without loss of generality, we will set to zero and some positive number which is total wealth y divided equally among r rich members of the total population of n. Total wealth is some given level of mean wealth multiplied by the size of the economy or $y \equiv \underline{y}n$. Then the only unequal pairs in the population are the r wealth holders interacting with the $n-r$ individuals without wealth, with both members of the pair experiencing a wealth difference of y/r. So, we have $\Delta = r(n-r)\underline{y}n/r$ and equation 1.3 becomes

(4) $$G = \frac{\Delta}{n(n-1)}\frac{1}{\underline{y}} = \frac{r(n-r)\underline{y}n/r}{n(n-1)}\frac{1}{\underline{y}} = \frac{n-1}{n-r}$$

from which we confirm that irrespective of population size, G = 1 when one person has all the wealth and that inequality declines as wealth is redistributed so as to be shared among a larger number of rich, that is, increasing r.

We can also see that holding constant the number of the wealthy, r,

(5) $$\frac{dG}{dn} = \frac{1-G}{n-1} > 0 \text{ for } G < 1$$

So G increases as the propertyless class increases (holding constant the size of the wealthy class). In the limiting case where a single person owns all of the wealth, increasing the number of propertyless does not affect the G. It is this limiting case where $r = 1$ so $G = 1$ that Debraj labels the "absurd consequence" of measuring inequality by G. Finally, to focus on the edges of the network rather than the nodes, we calculate the fraction δ of all pairs in which the two have a different wealth level,

(6) $$\delta = \frac{r(n-r)}{n(n-1)/2} = \frac{2r}{n}G$$

from which we see that if the population is divided into classes of equal size, then G itself is the fraction of interactions in which the members of the pair have different wealth (because $2r/n = 1$).

Note also that if there is just one wealth owner and one wealthless person, all of the interactions are unequal $\delta = 2/n = 1$ (there is just a single interaction). But as the propertyless class increases in size the fraction of all interactions that are unequal falls. How does it come about, then, that G increases with n (for a fixed number of rich households)? The answer is that as n rises (with r fixed) mean wealth remains unchanged (by assumption), but the mean wealth of the rich increases, and so the wealth difference between members of an unequal pair (that is $\underline{y}n/r$) also increases.

6. Is invariance to population replication a desideratum for a measure of inequality?

An attractive feature of the now-common variant of the Gini coefficient (equation 1.1, G^L) is that (for populations of any size) it is equal to the area between the perfect equality line and the Lorenz curve divided by one-half . But as equation 1.1 makes clear, G^L includes in the measure of societal inequality the "equality" of one's own wealth with one's own wealth (what we term the "fictive zeros"), violating our "economic inequality is social" axiom (see Fig. 2). Because as the term is conventionally used, societal inequality concerns relationships between people and is meaningless in a "perfectly equal" one-person society (where $G^L = 0$), we find this aspect of G^L to be a reason not to use it where it differs appreciably from G.

The payoff to including the fictive self-on-self equality comparisons in the measure of societal inequality is said to be that, defined in this way, it conforms to another axiom ("population symmetry" in Rajiv's memo), namely, that the measure of inequality should be invariant to replication of its members, so that inequality would be the same (maximal concentration) in a two-person society in which one held all of the wealth, in a four-person society in which two people equally shared all of the wealth, or if one thousand equally shared the wealth with another thousand who were without wealth, and so on.

Thus, the counterintuitive fictive zeros are the price of invariance to population replication, an axiom that is thought to be sufficiently intuitively appealing to justify setting aside the problem of including the self-on-self comparisons in the inequality measure.

However, from the perspective of experienced inequality, invariance to population replication is a bug not a feature of equation 1.1. Let's think about three economies with 2, 4, and 6 people, in each of which, half of the population owns all of the wealth in equal shares, and as before, total wealth is proportional to population size. Is it sufficiently obvious that the three economies are "equally unequal" with the level of equality unaffected by population replication that violation of this axiom disqualifies an inequality measure?

In all three societies G is also equal to the fraction of all interactions that are between people of differing wealth levels (because $2r = n$, from equation 2.3, $\delta = G$). This falls from 100 percent when $n = 2$ to two thirds for $n = 4$ and to 3/5 for $n = 6$. The wealth difference between the members of the unequal pairs is unchanged as the population replicates (because n/r by design is unchanging). The only thing that has changed is that in the larger population people experience interactions with others of the same wealth level more commonly. Our intuition is that G correctly shows that experienced inequality in the larger societies is less, and that as a result, the invariance to population replication axiom is unappealing or at least not so intuitive that its violation would disqualify G.

To understand what the invariance to population replication means, think about how the population might replicate and the conditions under which inequality would be invariant. In the above example, let the number of rich and poor increase from 2 in each to 4 in each but now, let the 2 poor and the 2 rich form single households and so on. Then because there are just two households, one rich and the other poor, relative between household inequality would remain unchanged. The reason is that the additional within class pairings of equals – which leads G to fall as n and r both rise proportionally – would all be within these two fictive households and hence would be ignored in the population-invariant G^L measure of inequality. In the absence of the fictive formation of households or some equivalent device, inequality falls as the population is replicated due to the more than proportional increase in the number of pairings (edges in Fig. 1) with zero inequality.

Let's now reconsider what Debraj finds "even more absurd" about G, namely that a thousand-person society with just two wealth holders is more equal than a 2-person society with one person holding all the wealth. We agree that these extreme cases challenge the intuitions. In the large population, the vast majority of interactions are among people with the same wealth while in the latter none are. This does not settle which is "really" more unequal – where they exist, the wealth differences in the thousand-person society (500 times mean wealth rather than twice) are much greater than in the two-person society. But the example does suggest that it is far from absurd that the former would be perceived in some sense as more equal than the latter.

For more than one wealth holder, both G and G^L increase as the number of propertyless increases. For example, for a given $r = 2$, G is 0.5 for $n = 3$ and 0.95 for $n = 20$, while G^L is respectively 0.333 and 0.90. It would be helpful to clarify the intuition for inequality increasing with increasing numbers of the propertyless. Once clearly articulated, we could then assess whether the best way to capture that intuition is to count self-on-self comparisons as if they were real social relationships in which wealth inequality is entirely absent, or instead use a measure that captures the very same intuition, but does not count the fictive zeros.

The above discussion also recommends a network representation of inequality, if capturing the experience of inequality is an objective. We have motivated our examples by a random pairing environment equivalent to a complete network. But societies differ greatly in who interacts with (or even is aware of) whom. For example, if the social structure in question is a star with the wealth holder at the center, then the fraction of one's interactions that are with someone of a different wealth level (all of them, $\delta = 1$) is a constant as the size of the network grows, and as a result δ will be greater (for $n > 2$) than in the complete network that we have used above as our illustration.

Debraj's intuition is not incorrect (how could it be?) But it is far from obvious to us. And competing and quite different intuitions may also be appealing.

References

Paul D Allison. Reply to Jasso. *American Sociological Review*, 44(5):870–872, 1979.

Anthony Atkinson. *The Economics of Inequality*. Oxford University Press, 1983.

Samuel Bowles and Wendy Carlin. Inequality as experienced difference: A reformulation of the Gini coefficient. *Economics Letters*, 186:108789, 2020.

Lidia Ceriani and Paolo Verme. The origins of the Gini index: Extracts from Variabilità e Mutabilità (1912) by Corrado Gini. *The Journal of Economic Inequality*, 10(3):421–443, 2012.

Frank A Cowell and Fatemeh Mehta. The estimation and interpolation of inequality measures. *The Review of Economic Studies*, 49(2):273–290, 1982.

Partha Dasgupta, Amartya Sen, and David Starrett. Notes on the measurement of inequality. *Journal of Economic Theory*, 6:180–187, 1973.

Angus Deaton. *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy*. World Bank Publications, 1997.

Joan Esteban and Debraj Ray. On the measurement of polarization. *Econometrica*, 62:819–851, 1994a.

Joan-Maria Esteban and Debraj Ray. On the measurement of polarization. *Econometrica: Journal of the Econometric Society*, pages 819–851, 1994b.

Gary Fields and John Fei. On inequality comparisons. *Econometrica*, 46:306–316, 1978.

James Foster. An axiomatic characterization of the Theil measure of income inequality. *Journal of Economic Theory*, 31:105–121, 1983.

Joseph L Gastwirth and Marcia Glauberman. The interpolation of the Lorenz curve and Gini index from grouped data. *Econometrica*, 44(3):479–483, 1976.

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

Corrado Gini. Sulla misura della concentrazione e della variabilità dei caratteri. *Atti del Reale Istituto Veneto di Scienze, Lettere ed Arti*, 73:1203–1248, 1914.

Corrado Gini. On the measurement of concentration and variability of characters. *International Journal of Statistics*, 63(1):1–38, 2005.

Guillermina Jasso. On Gini's mean difference and Gini's index of concentration. *American Sociological Review*, 44(5):867–870, 1979.

Max O Lorenz. Methods of measuring the concentration of wealth. *Publications of the American statistical association*, 9(70):209–219, 1905.

Debraj Ray. *Development Economics*. Princeton University Press, 1998.

Michael Rothschild and Joseph Stiglitz. Increasing risk: I. A definition. *Journal of Economic Theory*, 2:225–243, 1970.

Amartya Sen. *On Economic Inequality*. Oxford University Press, 1973.

S. Subramanian. An elementary interpretation of the Gini inequality index. *Theory and Decision*, pages 375–379, 2002.

S. Subramanian. Variable populations and inequality-sensitive ethical judgments. In Banasri Basu, Satya R. Chakravarty, Bikas K. Chakrabarti, and Kausik Gangopadhyay, editors, *Econophysics and Economics of Games, Social Choices and Quantitative Techniques*, pages 181–191. Springer Milan, 2010.

Dominique Thon. An axiomatization of the Gini coefficient. *Mathematical Social Sciences*, 2(2): 131–143, 1982.

Shlomo Yitzhaki and Edna Schechtman. *The Gini methodology: A primer on a statistical methodology*. Springer, 2013.