2021

2021-1 Personal Gini Coefficients

James B. Davies

Personal Gini Coefficients

by

James B. Davies

Research Report # 2021-1 **March 2021**

Department of Economics

Research Report Series

Department of Economics
Social Science Centre
Western University
London, Ontario, N6A 5C2
Canada

Personal Gini Coefficients

James B. Davies

University of Western Ontario

March, 2021

Abstract

The Gini coefficient is based on the sum of pairwise income differences, which can be decomposed into separate sums for individuals. Differences vis-à-vis poorer people represent an individual's *advantage*, while those with respect to richer people constitute *deprivation*. Weighting deprivation and advantage differently produces a family of personal Gini coefficients whose population averages each equal the overall Gini coefficient. Properties of the personal indexes explain why the Gini coefficient is most sensitive to changes in the middle of typical income distributions. Behavior of the personal indexes also throws light on the inequality impacts of secular changes in income distribution. In a simple Kuznets-type process, the Gini coefficient first rises and then falls but, throughout, a personal Gini coefficient will be rising for people in the traditional sector, while it is falling for those in the modern sector. In a leading case, the population shifts associated with polarization in labor markets in advanced economies also reduce personal inequality at the top and increase it at the bottom. The shift of population toward the two extremes unambiguously raises personal inequality for those in the middle. The wage changes accompanying polarization can, however, reverse these results, particularly at the top, as illustrated by calculations for U.S. polarization between 1980 and 2005.

Contact:
James Davies
jdavies@uwo.ca
Department of Economics
University of Western Ontario

I. Introduction

The Gini coefficient has a natural interpretation as the mean of personal inequality assessments. While that fact is obvious once pointed out, it was not emphasized in the original work by Gini (1914) and has not been highlighted since. This paper shows that this straightforward interpretation throws important light on the properties of the Gini coefficient. It also allows us to better understand the reaction of the Gini coefficient, and possibly also that of individuals, to secular changes in income distribution. The latter include the hypothetical transition from a traditional to a modern economy analyzed by Kuznets (1955), and the polarization in income distribution seen in recent decades in the US and some other countries. Personal assessments of the direction of change in inequality may differ between people at different income levels. These results suggest that our understanding of inequality measurement can be enriched by studying what it may mean at the personal level.

The Gini coefficient can be defined or interpreted in many ways (Yitzhaki, 1998). For our purposes the most useful is that it equals one half the mean difference divided by the mean. That is, the Gini coefficient can be found by taking the sum of all pairwise absolute income differences, S, converting to an average and normalizing by the mean income. S equals the sum across individuals $i = 1, .., n$ of their personal sums of income differences with all other individuals, S_i. The latter provide the basis for a personal inequality index whose average across the population is the Gini coefficient.

For each individual, S_i is composed of the sum of differences with respect to people at higher incomes plus the sum of differences vis- à -vis lower income persons. Following Yitzhaki (1979) the sum of differences with higher incomes may be used to define the individual's *deprivation*. That concept is complemented by the individual's *advantage*, derived from the sum of differences with respect to lower incomes.[1] Summing either deprivation or advantage across the whole population produces the same total

[1] Yitzhaki (1979) used the term "relative deprivation", which was introduced by Runciman (1966) to refer to any case in which some members of a reference group felt deprived compared to other members of their group.

(Yitzhaki, 1979). Therefore, any weighted average of deprivation and advantage, as well as an unweighted average, will generate the Gini coefficient. This means that there is a whole family of personal Gini coefficients. One implication is that if societies choose to base overall inequality measurement on an average of individual assessments they may all use the Gini coefficient at the aggregate level even if they differ in the weight placed on advantage vs. deprivation at the personal level.

The personal inequality indexes discussed here have both "top down" and "bottom up" interpretations. This paper does not take a position on which viewpoint is preferable. It is not necessary to make a choice in order to pursue the analysis. A personal Gini coefficient could be regarded as reflecting how a social planner would measure personal inequality. This is a "top down" interpretation. The "bottom up" view is that it would be reasonable for individuals themselves to assess inequality using such a measure. Why might they do so? One possibility is that they could have interdependent utility functions of a form that suggests the use of a personal Gini coefficient (Fehr and Schmidt, 1999). Another may lie in bounded rationality. Holding the mean constant, in a society of two people the difference between their incomes is a natural indicator of inequality. People might, implicitly if not explicitly, extend this to regard the average of pairwise differences as an attractive indicator of inequality when there are more than two people. That conclusion could be reinforced by information and computational constraints. As shown in this paper, in order to compute the value of a personal Gini coefficient the individual only needs to know the fractions of the population with income above and below her and the average incomes of those two groups. While we should not suppose that real-world individuals know everyone else's income, they might be able to make a serviceable guess at these fractions and averages.

If the "bottom up" interpretation of personal Gini coefficients is taken, our analysis is clearly related to the literature on individual attitudes toward inequality. A portion of that literature attempts to measure

"Deprivation" is used here simply because it is shorter. Fehr and Schmidt (1999) referred to the same concept as "disadvantageous inequality", but the term deprivation still dominates in the literature. Yitzhaki (1979) used "satisfaction" rather than "advantage". "Advantage" is used here as a more neutral term.

attitudes within narrow reference groups, e.g. co-workers or members of the same occupation. In such cases people tend to be averse to deprivation but to like advantage. As Clark and D'Ambrosio (2015) point out, in the income distribution literature the usual reference group is broader. In that context, following Yitzhaki (1979, 1982) and Fehr and Schmidt (1999) the general expectation has been that people will be averse to both deprivation and advantage. There are some empirical or experimental studies that have estimated aversion to deprivation and/or advantage with broader reference groups. Using the German SOEP survey data, D'Ambrosio and Frick (2007) find strong aversion to deprivation (but do not report on attitudes to advantage). Cojocaru (2014) finds significant aversion to both advantage and deprivation using a survey of 27 transition countries. In experiments with subjects who played a sequential public goods game, Teyssier (2012) found that 40% were averse to both advantage and deprivation while 18% were averse to neither. While these studies do not provide strong evidence on the relative degree of aversion to deprivation vs. advantage, neural studies find that brain activity reacts more strongly to deprivation (Clark and D'Ambrosio, 2015) and there appears to be close to a consensus that aversion is stronger to deprivation than to advantage for most people.

The remainder of the paper proceeds as follows. For expositional simplicity we start by working with the "unbiased" case in which advantage and deprivation are equally weighted. Section II defines the unbiased personal Gini coefficient and derives some of its basic properties. In Section III we then explore how the behavior of this index helps to explain the sensitivity of the Gini coefficient to income changes in different ranges of a distribution. The analysis is extended to allow unequal weighting of deprivation and advantage in Section IV, which shows that the main insights of the previous two sections survive this generalization. How the personal assessments of inequality vary with income is discussed in Section V and the behavior of those assessments during periods of secular change in income distribution is examined in Section VI. Section VII concludes.

II. Unbiased Personal Gini Coefficient

In this section we see how the Gini coefficient can be defined as the average value across individuals of an "unbiased" personal Gini coefficient, and begin to examine its properties. The Gini coefficient for an income distribution equals one half the mean difference divided by the mean, as in:

$$(1) \qquad G = \frac{1}{2n^2\bar{y}} \sum_{i=1}^{n} \sum_{j=1}^{n} |y_i - y_j| = \frac{S}{2n^2\bar{y}}$$

where y_i is the income of individual i, $\bar{y}$ is mean income, $n > 1$, $y_1 \leq y_2 \leq \cdots \leq y_n$, S is the sum of differences, and S/n^2 is the mean difference.[2]

A natural but previously overlooked interpretation is that G is the mean value across individuals of an unbiased personal Gini coefficiet, G_i:

$$(2) \qquad G = \frac{1}{n} \sum_{i=1}^{n} G_i$$

where

$$(3) \qquad G_i = \frac{1}{2n\bar{y}} \sum_{j=1}^{n} |y_i - y_j| = \frac{S_i}{2n\bar{y}}$$

and S_i is the sum of differences for individual i. Equation (3) can be rewritten:

$$(4) \qquad G_i = \frac{1}{2n\bar{y}} \left[n_i^l \left(y_i - \bar{y}_i^l \right) + n_i^h \left(\bar{y}_i^h - y_i \right) \right]$$

[2] As mentioned earlier, the Gini coefficient can be expressed in many different ways (Yitzhaki, 1998). This is one of the two principal forms in which it was originally set out in Gini (1914), and is the most convenient for our discussion.

where n_i^l is the number of individuals with income less than or equal to y_i, excluding individual i, and n_i^h is the number with income strictly greater than y_i, so that $n_i^l + n_i^h = n - 1$.[3] $\bar{y}_i^l$ and $\bar{y}_i^h$ are mean income among those with income less than or equal to y_i, excluding i, and strictly greater than y_i respectively.

Let H_i be the set of all *j* such that $y_j > y_i$, and L_i be the set of all *j* excluding *i* such that $y_j \leq y_i$. Equation (4) can be expressed as:

$$(4') \qquad G_i = \frac{1}{2\bar{y}}(A_i + D_i)$$

where:

$$(5i) \qquad A_i = \frac{n_i^l}{n}\left(y_i - \bar{y}_i^l\right) = \frac{1}{n}\sum_{j \in L_i}(y_i - y_j)$$

$$(5ii) \qquad D_i = \frac{n_i^h}{n}\left(\bar{y}_i^h - y_i\right) = \frac{1}{n}\sum_{j \in H_i}(y_j - y_i)$$

D_i is the discrete analogue of the measure of relative deprivation for an individual, which we will refer to simply as deprivation, proposed by Yitzhaki (1979) for a continuous distribution. It equals the average shortfall of *i*'s income below the income of those who are better off, weighted by the fraction of the population in the latter group. Equation (4′) shows that G_i is the simple average of D_i and a complementary measure, A_i, normalized by the mean. We will say that A_i represents individual *i*'s *advantage* compared to people with lower income. Thus, from the individual perspective inequality consists of both deprivation with respect to the better off and advantage over the worse off.

While G_i is a natural personal inequality index to associate with the Gini coefficient, it is not the only such index. As mentioned earlier, and as shown in Section IV, one can define a more general class of

[3] The choice to include individuals who have the same income as *i* in the lower group rather than in the higher group is arbitrary but does not affect the results in any significant way.

personal Gini coefficients that are based on a weighted average of A_i and D_i. G_i is a special case in which the weights on A_i and D_i are equal.

From (4) we have:

Proposition 1: *G_i is insensitive to a transfer of income within H_i or within L_i if the composition of neither group changes as a result of the transfer.*

The proposition follows from the fact that transfers of income confined either to H_i or L_i do not alter n_i^l, $\bar{y}_i^l$, n_i^h, or $\bar{y}_i^h$ or any other term on the right-hand side of (4). In terms of (4′), as noted by Yitzhaki (1979) these transfers have no effect on advantage, A_i, or on deprivation, D_i. The insensitivity of G_i to such transfers means that it does not respect the Pigou-Dalton principle of transfers, which is a cornerstone of the theory of aggregate inequality measurement.[4] That an aggregate index that respects the Pigou-Dalton principle can be built on the basis of personal indexes that violate the principle is striking.

Sensitivity of G_i to a transfer of income between H_i and L_i

What determines how sensitive G_i is to a transfer of income between H_i and L_i? Consider the transfer of a total amount *R* from H_i to L_i. Note that such a transfer reduces both A_i and D_i by *R/n*, as can be seen from (5) where $n_i^l(y_i - \bar{y}_i^l)$ and $n_i^h(\bar{y}_i^h - y_i)$ both fall by *R*. We will allow *R* to be negative, so this also handles the case of transfers from L_i to H_i, which *increase* A_i and D_i by equal amounts. Using

$$\frac{\partial A_i}{\partial R} = \frac{\partial D_i}{\partial R} = \frac{-1}{n}$$

from (4′) we have:

[4] Dalton (1920, p. 351) identified the central importance of the "principle of transfers", which says that a rank-preserving transfer from a richer person to a poorer person reduces inequality. Dalton referred his readers to an earlier statement of the same idea by Pigou. More recently the principle has come to be referred to as the "Pigou-Dalton" principle of transfers (Sen, 1973).

$$(6)\quad \frac{\partial G_i}{\partial R} = -\frac{1}{n\bar{y}}$$

which allows us to state:

Proposition 2: When income is transferred from a person with income strictly above y_i to someone with income strictly below y_i, G_i falls, while if income is transferred from a person with income strictly below y_i to someone with income strictly above y_i, G_i rises. In both cases the change in G_i is proportional to the amount transferred and independent of y_i.

Sensitivity of G_i to a transfer affecting y_i

We also need to analyze those cases where distributional changes affect individual *i*'s own income. There are two situations to consider. One is that of a transfer from *i* to another person *j*. The other is that of a transfer from *j* to *i*. We will consider them in turn. *In this analysis, and in the remainder of the paper unless indicated otherwise, we will assume $y_1 < y_2 < \cdots < y_n$.* This assumption will simplify the analysis since, for example, it implies that when *n* is odd there is a unique individual with median income, y^{med}, and half the remaining population has $y_i < y^{med}$ while the other half have $y_i > y^{med}$.[5] If *n* is even there is no individual with $y_i = y^{med}$, but y^{med}, which is defined as the midpoint between $y_{n/2}$ and $y_{n/2+1}$, again divides the population into two sub-populations of equal size with incomes above and below the median.

Transfer from *i* to *j*: Let y_i^o and y_j^o be initial incomes and consider the effect on G_i of the transfer of a small amount *r* from individual *i* to individual *j*. From (4) we obtain:

Proposition 3a: The effect on G_i of a small transfer in the amount of *r* from individual *i* to an individual *j* is given by:

[5] If we assume only $y_1 \leq y_2 \leq \cdots \leq y_n$ then there could be multiple individuals with median income and the groups with income strictly below the median and strictly above the median need not contain an equal number of members. Consider for example a population with the set of incomes (1, 1, 2, 2, 2, 3).

$$(7i) \qquad \Delta G_i = \frac{1}{2n\bar{y}}\left[\left(n_i^h - n_i^l\right) - 1\right]\mathrm{r}\,, \qquad i > j$$

$$(7ii) \qquad \Delta G_i = \frac{1}{2n\bar{y}}\left[\left(n_i^h - n_i^l\right) + 1\right]\mathrm{r}\,, \qquad i < j$$

If we would ignore the -1 and +1 in the square brackets on the right-hand side, (7) would say that irrespective of whether *i* was greater or less than *j*, a transfer from *i* to anyone else would increase G_i if *i* was below the median and reduce G_i if *i* was above the median. This reflects the fact that the main impact of the transfer on G_i is to reduce A_i and increase D_i. If $n_i^h > n_i^l$, individual *i* is below the median and from (5) we see that the increase in D_i will exceed the drop in A_i, since those changes are proportional to n_i^h and n_i^l respectively. If $n_i^h < n_i^l$, individual *i* is above the median and we have the opposite case. These conclusions are modified only trivially by the -1 and +1 in the square brackets in (7).[6]

<u>Transfer from *j* to *i*:</u> Here incomes after a transfer are $y_i^o + r$ and $y_j^o - r$. and we have:

Proposition 3b: The effect on G_i of a small transfer in the amount of *r* from an individual *j* to individual *i* is given by:

$$(8i) \qquad \Delta G_i = \frac{1}{2n\bar{y}}\left[\left(n_i^l - n_i^h\right) + 1\right]\mathrm{r}\,, \qquad i > j$$

$$(8ii) \qquad \Delta G_i = \frac{1}{2n\bar{y}}\left[\left(n_i^l - n_i^h\right) - 1\right]\mathrm{r}\,, \qquad i < j$$

[6] The -1 in (7i) means that the rank at which ΔG_i switches from being positive to negative as we go up the income scale in the $i > j$ case is shifted one position higher because the transfer goes to a person with income lower than the "donor" *i*, reducing $\bar{y}_i^l$ and A_i a little. And the +1 in (7ii) means that when $i < j$, ΔG_i switches from positive to negative is shifted one position *lower* than would otherwise be the case since the transfer goes to a higher income person, raising $\bar{y}_i^h$ and D_i a little.

Now the main effect of the transfer is to raise y_i and therefore to increase A_i and reduce D_i, which is equalizing if y_i is below the median and disequalizing if y_i is above the median. Again the point at which ΔG_i switches sign as i rises is offset one position by the small impact of the change in y_j on A_i when $i > j$ and on D_i when $i < j$.

Summing up, we can say that for an individual whose income is above the median, a small transfer from herself to someone else is equalizing, from a personal standpoint, if her income is above the median, and is disequalizing if her income is below the median (subject to the small qualification indicated in footnote 6). If she is the *recipient*, a small transfer is equalizing from the personal viewpoint if she is below the median and disequalizing if she is above the median (again subject to footnote 6). Thus, the situation in this form of personal inequality measurement is quite different from that in aggregate inequality measurement. In the latter, the impact of a small transfer on inequality is deemed equalizing if the donor's income exceeds the recipient's and disequalizing if the opposite holds. In the case of personal Gini coefficients, in contrast, whether the transfer is considered equalizing or disequalizing depends almost solely on the income of the person for whom the assessment is being made. For low income people, if they make a transfer it is disequalizing while if they receive a transfer it is equalizing. For high income people the opposite holds.

III. Explaining the sensitivity of the Gini coefficient to changes in different ranges of the income distribution

From (1) one may derive:

(9) $$G = \frac{2}{n^2\bar{y}}[y_1 + 2y_2 + 3y_3 + \cdots + ny_n] - \frac{n+1}{n}$$

(see e.g. Cowell, 2011, p. 114). This provides insight into the sensitivity of the Gini coefficient to changes in different ranges of the income distribution. Consider a small transfer, *r*, from individual *j* to individual *i* where $i < j$. This is an example of what would be called an "equalizing transfer" in discussions of aggregate inequality. From (9), this transfer will produce a change in the Gini coefficient given by:

(10) $$\Delta G = \frac{-2r(j-i)}{n^2\bar{y}}$$

which also tells us the impact of a transfer from *i* to *j*, in which case $r < 0$. We see that the impact on the Gini coefficient does not depend on y_i or y_j, but varies only with *r* and the difference in income ranks between *i* and *j*.

The fact that the sensitivity of the Gini coefficient to transfers is independent of the incomes of the transferor and transferee, but depends on the number of people between them in the distribution, is one of the most interesting properties of the Gini coefficient. This property follows directly from those of the personal inequality index G_i captured in Propositions 1, 2 and 3 above. Again considering a small transfer, *r*, from individual *j* to individual *i* where $i < j$, Proposition 1 implies:

(11i) $$\Delta G_k = 0. \quad k < i, k > j.$$

From Proposition 2 we have:

(11ii) $$\Delta G_k = \frac{-r}{n\bar{y}}. \quad i < k < j.$$

And from Proposition 3

(12) $$\Delta G_i = \frac{(n_i^l - n_i^h - 1)r}{2n\bar{y}}. \qquad \Delta G_j = \frac{(n_j^h - n_j^l - 1)r}{2n\bar{y}} \ .$$

Now, from (2) and (11i), the change in *G* resulting from a transfer from *j* to *i* is given by:

(13) $$\Delta G = \frac{1}{n}(\Delta G_i + \Delta G_j + \sum_{k=i+1}^{j-1} \Delta G_k)$$

Note first from (11*ii*) that

(14) $$\sum_{k=i+1}^{j-1} \Delta G_k = -(j-i-1)\frac{r}{n\bar{y}}$$

which is proportional to the number of people between *i* and *j*, that is the number of people the transfer from *j* to *i* "passes over".

Next, to complete the analysis of ΔG, note from (12) that:

$$\Delta G_i + \Delta G_j = \frac{(n_i^l - n_i^h - 1)r}{2n\bar{y}} + \frac{(n_j^h - n_j^l - 1)r}{2n\bar{y}}$$

$$= \frac{-r}{2n\bar{y}}[(n_j^l - n_i^l) + (n_i^h - n_j^h) + 2]$$

Since $n_j^l - n_i^l$ and $n_i^h - n_j^h$ both equal *j* – *i* we have:

(15) $$\Delta G_i + \Delta G_j = \frac{-r}{n\bar{y}}(j-i+1)$$

Hence, like $\sum_{k=i+1}^{j-1} \Delta G_k$, $\Delta G_i + \Delta G_j$ is proportional to the size of the transfer and rises linearly, in absolute value, with the number of people between *i* and *j*.[7] In this case the reason for dependence on the number of people between *i* and *j* is that the effects of the transfer cancel out for A_i and A_j on the one hand, and for D_i and D_j on the other, where the sums they are based on overlap. The range of overlap includes all $k < i$ for A_i and A_j, and all $k > j$ for D_i and D_j. The range where effects do not cancel out has $j - i + 1$ people in it.

Summing up, substituting (14) and (15) into (13) we have:

(16) $$\Delta G = \frac{-r}{n^2\bar{y}}[(j-i+1) + (j-i-1)] = \frac{-2r(j-i)}{n^2\bar{y}}$$

[7] Note that the right-hand-side of (15) is not *proportional* to the number of people between *i* and *j*, which is $j - i - 1$.

which is the same as (10). So, we have shown that the mean of the effects on the personal inequality indexes resulting from the transfer equals the change in *G* that one would expect from aggregate inequality analysis.

The purpose of this exercise has been to show that the effects of a transfer on personal inequality explain the impact on *G*. That the reaction of *G* is governed by the number of people between transferor *j* and transferee *i* is due to two things: (i) aside from *i* and *j* themselves, the only people whose personal inequality is affected by the transfer are the individuals between them in the distribution, and (ii) the effects of the transfer on G_i and G_j cancel out except for those based on changes in income gaps between *i* or *j* and individuals in the range (*i*+1, *j*-1).

IV. Unequal Weighting of Deprivation and Advantage

Yitzhaki (1979) defined relative deprivation for a society as a whole, *D*, as the average of individual deprivation indexes D_i. He worked with continuous distributions. The corresponding relationship with a discrete income distribution is:

(17) $$D = \frac{1}{n}\sum_{i=1}^{n} D_i$$

We can define overall advantage in a parallel way as:

(18) $$A = \frac{1}{n}\sum_{i=1}^{n} A_i$$

Yitzhaki shows that *D* is related to the Gini coefficient according to:

(19) $$G = \frac{D}{\bar{y}}$$

This result might appear puzzling, given that, from (4′), D_i represents only part of an individual's contribution to G_i and therefore to *G*. The explanation is as follows. The Gini coefficient is proportional to the sum of differences, *S*. We can arrange the pairwise differences $|y_i - y_j|$ making up *S* in a matrix *M* with *i* indexing rows and *j* indexing columns. *D* is the mean of the elements of *M* above the main diagonal while *A* is the mean of the below-diagonal elements. Now, the above-diagonal elements have the same mean as the below-diagonal elements in *M*, since e.g. $|y_2 - y_1| = |y_1 - y_2|$. Hence *A* = *D*. To get from *D* to *S* we must therefore double *D* and multiply by n^2 (to go from an average to a sum). The same procedure could be used to generate *S* from *A*. Thus we have $S = 2n^2 D = 2n^2 A$ or:

(20) $$A = D = \frac{S}{2n^2}$$

Substituting the expression for *D* from (20) into (19) we obtain $G = S/(2n^2\bar{y})$, that is equation (1).

While Yitzhaki's approach and ours are closely related, his D_i and our G_i are distinct. G_i depends not just on deprivation, D_i, but also on advantage, A_i. While, overall, *A* = *D*, at the individual level there is no such relationship. A_i rises and D_i falls as we move up through the income distribution from y_1 to y_n, and they do so at rates that rise or fall depending on the shape of the particular income distribution being examined.

The fact that $A = D$ has important consequences for personal Gini coefficients. Using (19) and $A = D$, *G* may be found by taking a weighted average of *A* and *D*, as in:

(21) $$G = \frac{\lambda A + (1-\lambda)D}{\bar{y}} \qquad 0 \le \lambda \le 1$$

where we require the weights to be positive. This in turn reveals that there is a family of personal Gini coefficients of the form:

(22) $$G_i^{\lambda} = \frac{\lambda A_i + (1-\lambda)D_i}{\bar{y}} \qquad 0 \le \lambda \le 1$$

Hence, while λ may differ across societies, they can nevertheless agree on using G as an aggregate measure of inequality. In the continuous case this result could be generalized to allow λ to differ across individuals, as long as the distribution of λ was independent of individual income.

We may ask which of the results derived above for the λ = ½ case survive once $\lambda \neq$ ½ is allowed. Proposition 1, which says that the G_i are insensitive to transfers entirely within the H_i or L_i comparator groups, survives. The principle is not affected by re-weighting income differences with the H_i and L_i groups via λ≠ ½ . Proposition 2, which says that when income is transferred from those with income above (below) y_i to those with income below (above) y_i the fall (rise) in G_i is proportional to the total amount transferred, R, and is independent of y_i is also unaltered because we still have:

$$\frac{\partial A_i}{\partial R} = \frac{\partial D_i}{\partial R} = \frac{-1}{n}$$

and (6) survives unchanged because in the more general formulation, using (22) we have:

$$(6') \qquad \frac{\partial G_i^{\lambda}}{\partial R} = \frac{1}{\bar{y}}\left[\lambda \frac{\partial A_i}{\partial R} + (1-\lambda)\frac{\partial D_i}{\partial R}\right] = -\frac{1}{n\bar{y}}$$

Proposition 3 described the impact on G_i of making a small transfer from another person to individual i. Assuming $y_1 < y_2 < \cdots < y_n$, the conclusion in the λ = ½ case was that, except for a very small region around the median, a transfer from a higher income person would reduce G_i if $y_i$was below the median, and increase G_i if y_i was above the median. Converse results held if the transfer came from a lower income person. The critical role of the median arose because with λ = ½, advantage, A_i, and deprivation, D_i, are equally weighted. In general, the critical percentile is given by 1-λ. Thus, for example, if one placed half as much weight on A_i as on D_i, i.e. λ = 1/3, the critical percentile would be 2/3. That means that a small transfer from someone with higher income would be regarded as equalizing for or by almost everyone in the bottom two thirds of the population, but as disequalizing for almost all of those in the top third. This occurs because putting a higher weight on D_i increases the equalizing impact on G_i^λ from the fall in D_i caused by such a transfer.

V. Personal Inequality Assessments at Different Income Levels

This section examines how G_i^λ varies as y_i rises from y_1 to y_n. Results are provided for the general case where λ can take on any value in the interval [0,1], but specific conclusions for the case where λ = ½ are also noted.

How does G_i^λ change as we move up through the distribution of income? We continue to assume $y_1 < y_2 < \cdots < y_n$. As we go from individual *i* to *i*+1, the absolute income gaps in (3) or implicitly in (22) increase in value by $y_{i+1} - y_i$ for all *j* such that $y_j < y_i$, and the corresponding gaps for all *j* > *i* fall by the same amount. Hence we should expect that G_i^λ will initially decline as *i* rises from 1, since at the start there are more people with *j* > *i* than with *j* ≤ *i* , until some critical point is reached, beyond which G_i should begin to increase. Formally we have:

Proposition 4: If $y_1 < y_2 < \cdots < y_n$,

$$G_{i+1}^\lambda \underset{<}{\overset{>}{=}} G_i^\lambda \quad \text{as} \quad \frac{i}{n} \underset{<}{\overset{>}{=}} 1 - \lambda \, .$$

Proof: See Appendix.

Proposition 4 indicates that G_i^λ falls up to the $(1-\lambda)100$th percentile of the distribution and increases above that. As indicated above, this U-shaped pattern is based on the fact that moving from income y_i to income y_{i+1}increases the income gaps with lower income people and reduces those with higher income people by the same absolute amount. The relative impact of changes in the upper gaps compared with that of changes in the lower gaps is (1-λ)/λ. This means that G_i^λ will fall more rapidly starting from *i* = 1 if λ < ½, compared with the λ = ½ case, and less rapidly if λ > ½. Note that if $\lambda = \frac{1}{2}$, $G_i^\lambda = G_i$ falls up to the 50th percentile, that is up to the median, and rises thereafter.

We can also readily identify the value of G_i^λ at the bottom and top of the distribution (i = 1 and i = n), as well as the value of G_i^λ for the median individual, G_{med}^λ, if n is odd. We have:

<u>Proposition 5:</u> If $y_1 < y_2 < \cdots < y_n$,

(i) $G_1^\lambda = (1-\lambda)(1-\frac{y_1}{\bar{y}})$

(ii) if n is odd, $G_{med}^\lambda = \frac{n-1}{2n\bar{y}}[(1-\lambda)\bar{y}_{med}^h - \lambda\bar{y}_{med}^l]$; if n is even, G_{med}^λ is not defined,

(iii) $G_n^\lambda = \lambda(\frac{y_n}{\bar{y}} - 1)$

Proof: See Appendix.

Proposition 5 allows us to put upper bounds on G_1^λ and G_n^λ. If y_1 is non-negative, the highest possible value of G_1^λ is $1-\lambda$, which occurs when $y_1 = 0$. When deprivation and advantage are weighted equally, that is when $\lambda = \frac{1}{2}$, the maximum value is $\frac{1}{2}$. But the maximum value of G_1^λ ranges from 0, when λ = 1 and personal inequality depends only on advantage, to 1 when λ = 0 and it depends only on deprivation. In view of Proposition 4, these maxima also apply to all G_i^λ up to the $(1-\lambda)100$th percentile.[8] The upper bound on G_n^λ occurs when one individual has all the income and $y_n = n\bar{y}$. In that case $G_n^\lambda = \lambda(n-1)$, which is also an upper bound for all G_i^λ's above the $(1-\lambda)100$th percentile.

Part (ii) of the proposition is also interesting, in throwing light on the value of the personal inequality index for the "average person", that is on the value of G_{med}^λ. The latter is based on a weighted average of $\bar{y}_{med}^h$ *and* $\bar{y}_{med}^l$, with the weight on $\bar{y}_{med}^h$ falling with λ. In the focal case with $\lambda = 1/2$, we have:

---

[8] Note that with λ = 1, the (1 – λ)100th percentile = 0, so that G_i^λ has no falling range.

$$G_{med} = \frac{(n-1)}{4n\bar{y}}(\bar{y}_{med}^{h} - \bar{y}_{med}^{l})$$

Since in any real-world example $(n-1)/n \approx 1$, this says:

$$G_{med} \approx \frac{\bar{y}_{med}^{h} - \bar{y}_{med}^{l}}{4\bar{y}}$$

In the U.S. today, for household income before tax, $\bar{y}_{med}^{h} \approx \frac{8}{5}\bar{y}$ and $\bar{y}_{med}^{l} \approx \frac{2}{5}$, which yields $G_{med} \approx 0.3$, less than the value of the Gini coefficient, which was 0.476 in 2013.[9] We may also note values of G_{med} under some familiar continuous distributions. G_{med} would equal $\frac{1}{4}$ for a uniform distribution, and if $y_i \sim N(\mu, \sigma)$, it would equal $\frac{2}{5}\frac{\sigma}{\mu}$, that is two-fifths of the coefficient of variation.

We can see that G_i^{λ} will generally not be symmetric around the median. Looking at the $\lambda = 1/2$ case again, for example, G_i will never be greater than $1/2$ at the lowest income level, but can be very high at the top end. G_i is not bounded above by 1, unlike the Gini coefficient. $G_n = 1$ is reached when $\frac{y_n}{\bar{y}} = 3$. That ratio is exceeded in almost all real-world cases. This implies that, in a mathematical sense, from the standpoint of the rich there is more inequality than from that of the poor when $\lambda = 1/2$, which is intuitive. For the rich there are relatively few people whose incomes is close to theirs, meaning there is a large gulf between their income and most others'.

VI. Personal Inequality During Secular Change in Income Distribution

This section asks how G_i^{λ} can be predicted to behave at different income levels during periods of secular change in income distribution. We focus initially in each case on the $\lambda = 1/2$ case, in which deprivation and advantage are weighted equally, referring to $G_i^{1/2}$ simply as G_i, as above. We start with the Kuznets

[9] With the help of quintile share and other data from U.S. Census Bureau (2021) it can be estimated that $\bar{y}_{med}^{h} = 1.64\bar{y}$ and $\bar{y}_{med}^{l} = 0.36\bar{y}$.

transformation and go on to the polarization and rising inequality that has been seen in the U.S. and many other high income countries in the last few decades. The principles at work are explored with the help of illustrative examples.

Kuznets Transformation

Kuznets (1955) analyzed what may happen to income distribution and inequality in a growing economy where production is shifting from an initially large traditional agricultural sector to a modern sector. The modern sector eventually comprises most if not all of the economy. The consequences for inequality can be illustrated using a stylized model in which individual incomes are uniform within each of the sectors, higher in the modern sector, and unchanging during the growth process.[10] (The assumption that individual incomes do not change within the sectors is relaxed below.) In this case the Gini coefficient, G, rises until the fraction of the population in the modern sector, p, hits a critical value, after which it declines. This critical value of p is less than one half. That is because, while the mean difference has a maximum at $p = 1/2$, the mean, which appears in the denominator of the expression for G, is rising throughout, so G has already started to decline at $p = 1/2$.

The suggestion that countries should generally be expected to display an inverted U-shaped time profile of the Gini coefficient, that is the "Kuznets hypothesis", has been shown not to describe what has actually happened in many countries (Deininger and Squire, 1998; Frazer, 2006; Angeles, 2010). However, it is of theoretical interest and is relevant to the historical experience of some countries. A current example is China, where a vigorous movement of people from the countryside to urban areas with much higher average income has been going on since the onset of market reforms in the late 1970s. At the beginning of that transition the Gini coefficient fluctuated around 0.30 (Sicular, 2013). It rose to a peak of 0.437 in

[10] Kuznets considered a richer range of possibilities. He allowed unequal income distribution within both sectors and believed the leading case was one in which there was greater inequality in the modern sector than in the traditional, or agricultural, sector. He also considered the impacts of changes in the relative income, and of income inequality, in the modern vs. the agricultural sector over time. In most cases he found that as the relative population of the agricultural sector declined there was an initial increase in inequality followed by a decline.

2010 and then began to fall, reaching 0.385 by 2016 (World Bank, 2021). These trends could plausibly be due to a Kuznets process (Knight, 2014). [11]

The behavior of the Gini coefficient and unbiased personal Gini coefficients in rural and urban areas during the Kuznets transformation will be illustrated here using an example whose implications are shown in Figure 1. It is assumed that income of each person in the traditional sector is 12% of per capita income in the modern sector. This gap is sufficient for the peak value of *G* to be 0.5.

Taking the λ = ½ case to begin with, we will refer to the individual inequality measures of people in the low and high income groups as G_L and G_H respectively. Since no one is worse off than those in the low income group, $G_L = \frac{D_L}{\bar{y}}$, that is it is based entirely on deprivation, while $G_H = \frac{A_H}{\bar{y}}$ and is based wholly on advantage. Denoting income per person in the traditional sector y_L and in the modern sector y_H, and adopting corresponding notation for the number of persons in each sector, n_L and n_H, we have:

(23i) $$D_L = \frac{n_H}{n}(y_H - y_L)$$

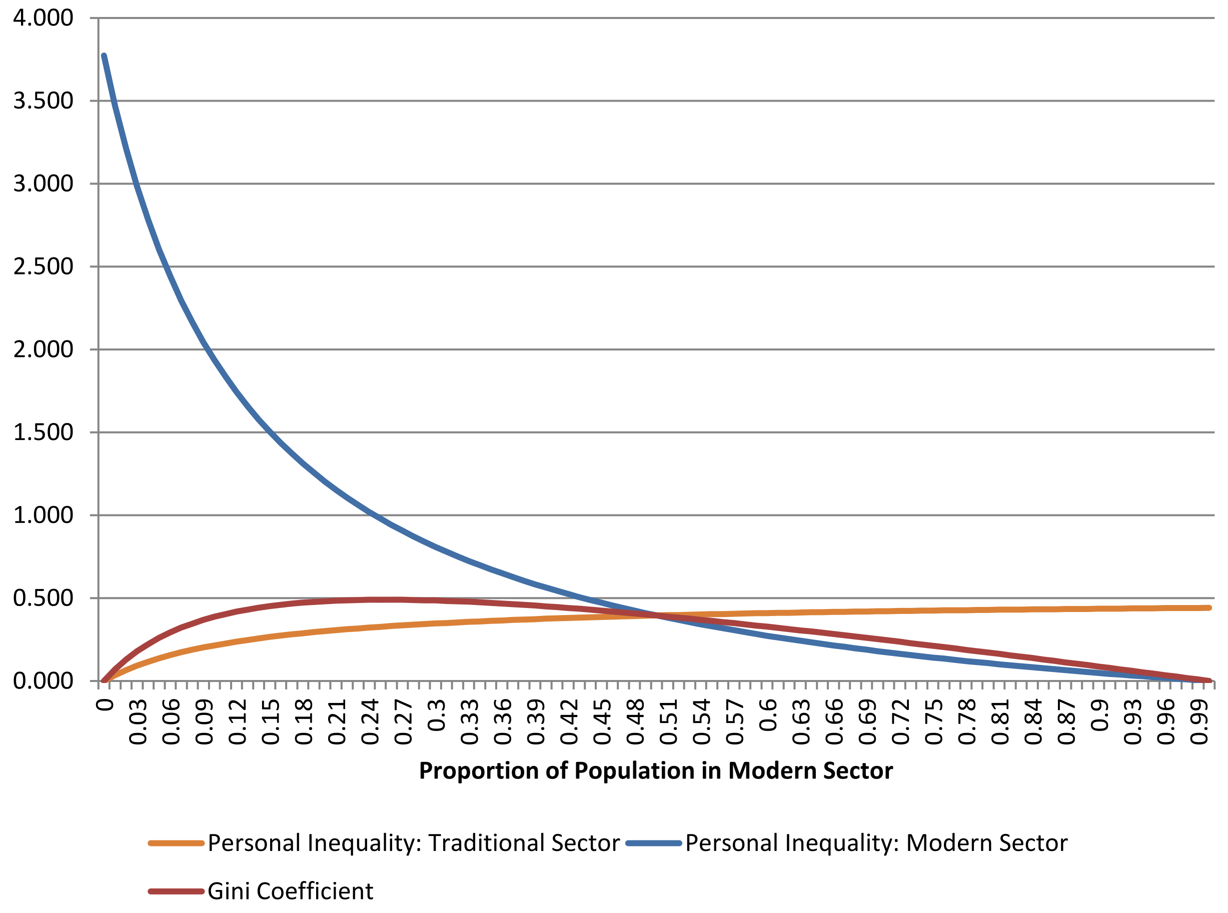
(23ii) $$A_H = \frac{n_L}{n}(y_H - y_L)$$

As shown in Figure 1, when the modern sector is tiny, G_L is close to zero. Almost everyone in the society has the same low income, so that n_H and D_L are very low. The situation in the modern sector is the opposite. Since almost everyone has much lower income than those in the modern sector, the individual inequality measure there, G_H is very high. Now, as development proceeds, G_L rises monotonically and G_H falls monotonically - - a necessary result in this simple model.[12] It is interesting to think what this

[11]Knight (2014) discussed whether China could be beyond the peak of the Kuznets curve. His conclusion was that that would depend in part on public policy but that there were strong underlying forces pushing in the direction of falling inequality in China.

[12] D_L rises with the increase in n_H and A_H falls as n_L declines. But one must also account for $\bar{y}$ rising throughout the Kuznets process when analyzing G_L and G_H. This rise reinforces the decline of A_H to ensure that G_H must fall throughout. And while the rise in $\bar{y}$, by itself, would make D_L fall, the increase in n_H has a stronger effect, so that G_L rises all through the process.

Figure 1: Kuznets Transformation: Personal Gini Coefficients with λ = 1/2 in Traditional and Modern Sectors with λ = 1/2 , and the overall Gini Coefficient



would mean if the personal inequality assessments reflected individual attitudes. People in the traditional sector would believe that inequality was becoming steadily worse while those in the modern sector would think the opposite, hardly a recipe for social harmony.

How does one resolve the conflict when the trend in inequality looks as radically different from the standpoint of two population groups as in the Kuznets transformation? The Gini coefficient offers a solution - - take an average of the personal assessments. Thus, in the Kuznets curve example, G is a population weighted average of the values of G_L and G_H. An alternative would be, in effect, to take a vote on the question of whether inequality was rising or falling - - a "democratic" approach. Here the democratic approach, based on personal inequality assessments, would say that inequality rises until p =

½ and falls thereafter. In the example, G says that inequality rises until p = ¼ and falls after that. That is because G_H falls faster than G_L rises, so that averaging G_H and G_L, even using population weights, places greater relative importance on the decline in G_H than on the rise in G_L.

The above analysis would not be affected significantly by moving from the λ = ½ case to the biased case with λ ≠ ½. There would of course be no impact on the time path of G. Since personal inequality in each sector only depends either on deprivation (in the traditional sector) or advantage (in the modern sector), at the individual level there would simply be a rescaling of G_L^λ and G_H^λ at each point in the Kuznets process. For a majority of people personal inequality would still be rising until p = ½ is reached, and above that point the opposite would still be true. G would have its peak at the same point as with λ = ½ . In terms of Figure 1, there would be a proportionate shift of the G_L curve by the factor $2(1-\lambda)$ and a shift of the G_H curve in the opposite direction by the factor 2λ. In the case where λ < ½, the G_L and G_H curves (now G_L^λ and G_H^λ) would move towards each other, while if λ > ½ the result would be the opposite.

What difference does it make if incomes are not constant within the two sectors during the Kuznets transformation? The question is whether changes in $(y_H - y_L)/\bar{y}$ can reverse those of $\frac{n_H}{n}$ or $\frac{n_L}{n}$ in the calculations of G_H^λ and G_L^λ, respectively. The answer depends on the percentage size of the possibly opposing changes. In the case of China, at least, the income changes appear to have been dominated by population shift. Identifying urban areas as our H sector and rural areas as L, from 1980 to 2014 $(y_H - y_L)/\bar{y}$ fell by 18% in China while $\frac{n_L}{n}$ dropped 44% and $\frac{n_H}{n}$ went up 183%.[13]

[13] The % changes reported here were calculated using tables 2-1 and 6-6 of National Bureau of Statistics of China (2016), which indicate an urbanization rate of 19.4% in 1980 and 54.8% in 2014. Disposable household income per capita in urban areas was 478 yuan in 1980 and 29,381 yuan in 2014; the corresponding rural numbers were 191 and 989 yuan.

Polarization

There is much theoretical and empirical literature on polarization (including Foster and Wolfson, 1992; Esteban and Ray, 1994; Acemoglu and Autor, 2011; Autor and Dorn, 2013; Green and Sand, 2015). Polarization in labor markets has received particular attention in the US and other high income countries in recent years. In this case the relative demand for labor shifts away from mid-level occupations to both low-skilled and (especially) high skilled occupations. Other things constant this should result in a shift in labor force composition away from the middle toward both the top and bottom. Such a shift has indeed occurred over significant timespans in the US, Canada, the UK, Germany and some other European countries (Acemoglu and Autor, 2011; Green and Sand, 2015). In most cases the wages of highly skilled workers have increased while those of workers in mid-level occupations have tended to decline. In the US it has also been found that wages have risen in certain low skilled occupations (Autor and Dorn, 2013) although that trend has not been seen in some other leading OECD countries (Green and Sand, 2015).

We will analyze the effects of labor market polarization on personal inequality in two steps, first considering only the effects of *population shift*, that is a rise in the number of individuals at low and high incomes combined with a reduction in the number at middle income. Subsequently we will look at the effect of changes in income. Assume that there are just three income levels in a society and that they display $y_L < y_M < y_H$. Numbers of individuals in the three groups are $n_L, n_M,$ and n_H. As in the Kuznets case the personal Gini coefficients of people in the bottom group and top groups are given by $G_L^\lambda = \frac{\lambda D_L}{\bar{y}}$ and $G_H^\lambda = \frac{(1-\lambda)A_H}{\bar{y}}$.

Once again, qualitative results for G_L^λ and G_H^λ will be the same as for G_L and G_H, so we will focus on the latter for simplicity. The increase in n_H due to polarization will tend to make A_H and G_H decrease since (from 5i):

(24) $$A_H = \frac{(n_L+n_M)}{n}(y_H - \bar{y}_H^l)$$

However, there is now an offsetting effect because $\bar{y}_H^l$ falls due to the population shift from the middle to lower groups, and therefore $(y_H - \bar{y}_H^l)$ increases. It can readily be shown that:

$$(25) \qquad \Delta A_H, \Delta G_H \overset{>}{\underset{<}{=}} 0 \quad \text{as} \quad \frac{\Delta n_L}{-\Delta n_M} \overset{>}{\underset{<}{=}} \frac{y_H - y_M}{y_H - y_L}$$

Now $\frac{y_H-y_M}{y_H-y_L} < 1$ and $\frac{\Delta n_L}{-\Delta N_M} < 1$ as well, so it is not immediately clear which way the inequality will go. However, we can make a prediction in a "leading case". With a positively skewed distribution of income we would have $\frac{y_H-y_M}{y_H-y_L} > \frac{1}{2}$, so that if half or fewer of those leaving the middle income group go to the lower group (which is in line with the experience in the US at least) we have $\frac{\Delta n_L}{-\Delta n_M} < \frac{1}{2}$ and A_H and G_H will decline, as in the Kuznets case.

Turning to the bottom group, from (5ii) we have:

$$(26) \qquad D_L = \frac{(n_M + n_H)}{n}(\bar{y}_L^h - y_L)$$

And it can be shown that:

$$(27) \qquad \Delta D_L, \Delta G_L \overset{>}{\underset{<}{=}} 0 \quad \text{as} \quad \frac{-\Delta n_M}{\Delta n_H} \overset{<}{\underset{>}{=}} \frac{y_H - y_L}{y_M - y_L}$$

Now, $\frac{y_H-y_L}{y_M-y_L} > 1$ and $\frac{-\Delta n_M}{\Delta n_H} > 1$ as well, so again there is ambiguity. However, in the leading case identified above $\frac{y_H-y_L}{y_M-y_L} > 2$ and $\frac{-\Delta n_M}{\Delta n_H} < 2$, so D_L and G_L will rise, as in the Kuznets analysis. This would be the result of the increase in $\bar{y}_L^h$ having a larger effect on D_L and G_L than the decline in $n_L^h = (n_M + n_H)$.

Population shift has unambiguous results for the middle group because, unlike the case of the top and bottom groups, the income differences vis-à-vis higher or lower groups are not affected by changes in $n_L, n_M,$ and n_H. G_M^λ depends on both A_M and D_M,

(28)
$$G_M^\lambda = \frac{\lambda A_M + (1-\lambda) D_M}{\bar{y}}$$

while personal advantage and deprivation are proportional to n_L and n_H respectively:

(29i)
$$A_M = \frac{n_L}{n}(y_M - y_L)$$

(29ii)
$$D_M = \frac{n_H}{n}(y_H - y_M)$$

Since both n_L and n_H rise in polarization, the pure effect of population shift is for G_M^λ to increase for any value of λ.

Turning to income changes, as mentioned above, it is typically observed in labor market polarization that y_H rises and y_M declines. In the US it has also been found that y_L rises. The rise of y_H opposes the "leading case" effects of population shift found above for the *H* group, so that personal inequality may *rise* at the top once income changes are taken into account. Impacts for both middle and lower groups are theoretically ambiguous. With $(y_H - y_M)$ rising, D_M also rises, tending to make G_M^λ increase. But A_M may fall if $(y_M - y_L)$ declines sufficiently. (The example considered below shows this can occur in practice.)   If A_M falls then G_M^λ will also fall if the weight placed on A_M in (28) is sufficiently large. Finally, the impact of income changes on G_L^λ is ambiguous since $(\bar{y}_L^h - y_L)$ in (26) may fall if y_L rises sufficiently and also because the fall of y_M reduces the change in $\bar{y}_L^h$, possibly even making it negative.

Given the theoretical ambiguity of the behavior of G_L^λ, G_M^λ and G_H^λ it is helpful to consider a real-world example. Autor and Dorn (2013) set out the changes in employment shares and wage rates for six broad occupational groups in the U.S. from 1980 to 2005. From Table 1, the top group, consisting of managers,

Table 1

Advantage A_i, Deprivation D_i, and Personal Gini Coefficients G_i with λ = 1/2, by Occupation Group - - Polarization Example Based on US Data, 1980 and 2005

Year	Occupation Group	Employment Share	Mean Wage (2004 $s)	A_i	D_i	G_i
1980	**1**	0.316	17.0	3.415	0	0.126
	2	0.048	15.6	2.524	0.412	0.108
	3	0.216	13.6	1.224	1.156	0.088
	4	0.099	11.9	0.528	2.117	0.098
	5	0.222	11.3	0.305	2.589	0.107
	6	0.099	8.2	0	5.364	0.198
2005	**1**	0.409	23.1	6.108	0	0.180
	2	0.030	15.2	1.425	3.241	0.137
	3	0.182	13.9	0.692	3.814	0.133
	4	0.046	12.7	0.399	4.716	0.150
	5	0.204	13.5	0.537	4.069	0.135
	6	0.129	9.6	0	7.413	0.218

Notes: (i) The mean wage is the geometric mean hourly wage derived from the mean log hourly wage reported by Autor and Dorn (2013), (ii) G_i is the personal inequality index G_i^λ when λ = 1/2**,** (iii) the occupational groups are:

1. managers, professionals, technicians, finance and public safety occupations.
2. production and craft occupations
3. transportation, construction, mechanics, mining and farm occupations
4. machine operators and assemblers
5. clerical and retail sales occupations
6. service occupations.

Source: Employment share and mean wage are from Autor and Dorn (2013, Table 1) - - see Note (i). The other columns were calculated by the author.

professionals, technicians, finance and public safety occupations experienced a 29% increase in employment share and a 36% rise in wage rates over those years. The middle four groups together had a 22% drop in employment share and only a 9% increase in wages. The bottom group, consisting of service occupations, had a 30% rise in employment share and a 17% increase in wages. These changes provide a dramatic example of labor market polarization.

For the sake of this example, assume that everyone within each of the six Autor and Dorn occupational groups has the same income. Using that assumption Table 1 shows G_i rising for all six groups, as does deprivation D_i (except for the top group, where it is identically zero.) On the other hand, advantage, A_i, falls for groups 2, 3 and 4, because their wages decline relative to the wage of the bottom group.[14] These results are obtained with $\lambda = 1/2$. If A_i is weighted sufficiently more heavily, G_i^{λ} declines from 1980 to 2005 for each of groups 2, 3 and 4. The critical values of λ are 0.61, 0.73 and 0.89 for groups 2, 3 and 4 respectively. Thus, if these groups were sufficiently more concerned about advantage than deprivation they would regard polarization as having reduced inequality between 1980 and 2005. While worth noting, this result may not affect one's conclusions much in view of the dominant opinion in the literature that $\lambda \leq ½$ likely holds for most people.

VII. Discussion and Conclusion

The fact that the Gini coefficient can be interpreted as the average of personal Gini indexes produces interesting insights. One important feature is that personal Gini coefficients are completely insensitive to transfers of income that occur only among people who have incomes above those of the reference individual, or among those with incomes below. This means that they do not obey the Pigou-Dalton principle of transfers. But they do regard transfers from those in the group above the individual to those

[14] Note that in 2005 group 4 has a lower mean wage than group 5. This is accounted for in the numbers shown in Table 1. In 2005 group 4 only has an advantage over group 6, while group 5 has an advantage over both groups 4 and 6.

in the group below as equalizing and transfers in the other direction as disequalizing. These properties explain why the Gini coefficient's sensitivity to transfers depends critically on the number of people with incomes between those of the donor and recipient, which makes it most sensitive to changes in the middle of typical income distributions. This dependence reflects the fact that, aside from the donor and recipient themselves, it is only people between them whose personal Gini coefficients are affected by the transfer.

As we have seen, each personal Gini coefficient is a weighted average of an individual's deprivation and advantage. That the relative weights placed on these components can vary has a range of implications. For example, the weights could vary across societies, perhaps reflecting differences in individual attitudes. At one extreme all the weight could be placed on deprivation in a society where people resented others being better off than themselves but had no concern about the income of those below them. At the other extreme all the weight could be put on advantage if everyone had been taught to have concern for the "less fortunate" and not to envy the better-off. And, of course, any weighting between these extremes could occur. But in each society, taking the average of personal Gini coefficient values would still yield the conventional Gini coefficient, since it is unaffected by the relative weight placed on deprivation vs. advantage. So, societies with quite different views about inequality at the individual level, could still all embrace the Gini coefficient as their aggregate measure of inequality. It is tempting to imagine that this might help to explain the wide international popularity of this index.

We have also discussed how personal inequality assessments may behave during secular change in income distribution. In the development context, in the simplest model of the Kuznets transformation, personal inequality for those in the traditional sector rises *throughout*, while the opposite occurs in the modern sector. If personal inequality reflects individual attitudes, the resulting scope for misunderstanding and conflict seems large. This may throw some light on the tensions that are observed during periods of rapid modernization. A further insight comes from the fact that the Gini coefficient says the Kuznets process stops being disequalizing well before half the population is in the modern sector. Thus, the direction of change in the Gini coefficient may not always reflect majority opinion.

Under polarization, population shifts not only to the top but also to the bottom, with a shrinking middle group. Income tends to rise at the top, fall in the middle, and may rise little at the bottom. In a leading case, population shifts increase personal inequality at the bottom and reduce it at the top, echoing the Kuznets transformation results. Personal inequality rises in the middle if there are no income changes. When income changes are also taken into account it is theoretically possible for any of the population shift effects to be reversed. Given this ambiguity we turned to the real world for some guidance. In an example based on the polarization seen in the US between 1980 and 2005, personal inequality rose for all groups when advantage and deprivation were equally weighted. For three middle groups personal inequality would have fallen if sufficiently more weight were placed on their advantage rather than their deprivation. However, if personal inequality reflects individual attitudes, it seems unlikely that the middle groups would indeed have regarded inequality as falling, given the broad consensus in the literature that most people tend to be more concerned about deprivation than advantage.

References

Acemoglu, D. and D.H. Autor (2011), "Skills, tasks and technologies: Implications for employment and earnings", *Handbook of Labor Economics* 4: 1043-1171.

Autor, David H. and David Dorn (2013), "The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market", *American Economic Review* 103 (5): 1553-1597.

Clark, Andrew E. and Conchita D'Ambrosio (2015), "Attitudes to Income Inequality: Experimental and Survey Evidence", chapter 13 in Anthony B. Atkinson and François Bourguignon, *Handbook of Income Distribution Vol. 2a*, 1st edition, North-Holland Elsevier: Amsterdam, 1147-1208.

Cojocaru, A. (2014), "Fairness and inequality tolerance: evidence from the Life in Transition survey", *Journal of Comparative Economics* 42 (3): 590-608.

Cowell, Frank A. (2011), *Measuring Inequality*, 3rd edition, Oxford: Oxford University Press.

Dalton, Hugh (1920), "The Measurement of Inequality of Incomes", *Economic Journal* 30: 348-361.

D'Ambrosio, Conchita and J.R. Frick (2007), "Income Satisfaction and Relative Deprivation: an Empirical Link", *Social Indicators Research* 81: 497-519.

D'Ambrosio, Conchita and J.R. Frick (2012), "Individual Well-Being in a Dynamic Perspective", *Economica* 79: 284-302.

Esteban, J. and D. Ray (1994), "On the measurement of polarization", *Econometrica* 62: 819–851.

Fehr, Ernst and Klaus M. Schmidt (1999), "A Theory of Fairness, Competition and Cooperation", *Quarterly Journal of Economics* 114: 817-868.

Fehr, Ernst and Klaus M. Schmidt (2003), "Theories of Fairness and Reciprocity: Evidence and Economic Applications", in Dewatripoint, M. L.P. Hansen, and S.J. Turnovsky (eds.), *Advances in Economic Theory, Eighth World Congress of the Econometric Society,* vol. 1, Cambridge University Press: Cambridge: 208-257.

Foster, James E. and Michael C. Wolfson (1992), "Polarization and the decline of the middle class", mimeo, reprinted 2010 in *Journal of Economic Inequality* 8: 247-273.

Gini, Corrado (1914), "On the measurement and variability of characters", *METRON-International Journal of Statistics,* LXIII (Part II): 3-38.

Green, David A. and Benjamin M. Sand (2015), "Has the Canadian labour market polarized?", *Canadian Journal of Economics* 48 (2): 621-646.

Knight, John (2014), "Inequality in China: An Overview", *The World Bank Research Observer* 29 (1): 1-19.

Kuznets, Simon (1955), “Economic Growth and Income Inequality”, *American Economic Review* 45: 1-28.

Li, Shi and Terry Sicular (2014), “The Distribution of Household Income in China: Inequality, Poverty and Policies”, *The China Quarterly* 217 (March): 1-41.

National Bureau of Statistics of China (2016), *China Statistical Yearbook – 2015*, www.stats.gov.cn accessed March 20, 2021.

Runciman, W.G. (1966), *Relative Deprivation and Social Justice*, London: Routledge and Kegan Paul.

Sen, Amartya (1973), *On Economic Inequality*, London: Oxford University Press (Clarendon).

Sicular, Terry (2013), “The Challenge of High Inequality in China”, *Inequality in Focus* 2(2): 1-8. http://www.worldbank.org/content/dam/Worldbank/document/Poverty%20documents/Inequality-In-Focus-0813.pdf

Teyssier, Sabrina (2012), “Inequity and risk aversion in sequential public good games”, *Public Choice* 51 (1-2): 91-119.

Yitzhaki, Shlomo (1979), “Relative Deprivation and the Gini Coefficient”, *Quarterly Journal of Economics*, 93 (2): 321-324.

Yitzhaki, Shlomo (1982), “Relative Deprivation and Economic Welfare”. *European Economic Review* 17: 99-113.

Yitzhaki, Shlomo (1998), "More than a Dozen Alternative Ways of Spelling Gini", *Research on Economic Inequality* 8: 13–30.

U.S. Census Bureau (2021), “Historical Income Tables: Income Inequality”, http://www.census.gov/data/tables/time-series/demo/income-poverty/historical-income-inequality.html Accessed March 20, 2021.

Appendix

This appendix provides proofs of propositions 4 and 5.

<u>Proposition 4:</u> If $y_1 < y_2 < \cdots < y_n$,

$$G_{i+1}^{\lambda} \overset{>}{\underset{<}{=}} G_i^{\lambda} \quad \text{as} \quad \frac{i}{n} \overset{>}{\underset{<}{=}} 1 - \lambda \, .$$

Proof: From (4′), (5) and (22), and using the assumption that $y_1 < y_2 < \cdots < y_n$, $G_{i+1}^{\lambda} \overset{>}{\underset{<}{=}} G_i^{\lambda}$ as:

$$\text{(A1)} \qquad \lambda \sum_{j=1}^{i} \left(y_{i+1} - y_j\right) + (1-\lambda) \sum_{j=i+2}^{n} \left(y_j - y_{i+1}\right) \overset{>}{\underset{<}{=}} \lambda \sum_{j=1}^{i-1} \left(y_i - y_j\right) + (1-\lambda) \sum_{j=i+1}^{n} \left(y_j - y_i\right)$$

Now, the left-hand side of this expression can be written:

$$\text{(A2)} \qquad \lambda\left[\textstyle\sum_{j=1}^{i-1}\left(y_i - y_j\right) + i(y_{i+1} - y_i)\right] + (1-\lambda)\left[\textstyle\sum_{j=i+1}^{n}\left(y_j - y_i\right) + (n-i)(\,y_i - y_{i+1})\right]$$

Hence, (A1) simplifies to:

$$\lambda\, i(y_{i+1} - y_i) + (1-\lambda)(n-i)(y_i - y_{i+1}) \overset{>}{\underset{<}{=}} 0$$

which is equivalent to:

$$\lambda\, i - (1-\lambda)(n-i) \overset{>}{\underset{<}{=}} 0$$

which becomes:

$$\lambda n + i - n \overset{>}{\underset{<}{=}} 0$$

from which one readily derives the result that

$$G_{i+1}^{\lambda} \overset{>}{\underset{<}{=}} G_i^{\lambda} \quad \text{as} \quad \frac{i}{n} \overset{>}{\underset{<}{=}} 1 - \lambda$$

Proposition 5: If $y_1 < y_2 < \cdots < y_n$,

(i) $G_1^\lambda = (1-\lambda)(1-\frac{y_1}{\bar{y}})$

(ii) if n is odd, $G_{med}^\lambda = \frac{n-1}{2n\bar{y}}[(1-\lambda)\bar{y}_{med}^h - \lambda\bar{y}_{med}^l]$; if n is even, G_{med}^λ is not defined,

(iii) $G_n^\lambda = \lambda(\frac{y_n}{\bar{y}} - 1)$

Proof: (i) From (5) and (22), given that $n_1^l = 0$,

$$G_1^\lambda = \frac{(1-\lambda)}{n\bar{y}}[n_1^h(\bar{y}_1^h - y_1)]$$

$$= \frac{(1-\lambda)}{n\bar{y}}[\sum_{j=2}^{n} y_j - (n-1)y_1]$$

$$= \frac{(1-\lambda)}{n\bar{y}}(\sum_{j=1}^{n} y_j - ny_1)$$

$$= \frac{(1-\lambda)}{n\bar{y}}(n\bar{y} - ny_1)$$

$$= (1-\lambda)(1-\frac{y_1}{\bar{y}})$$

(ii) From (5) and (22), if n is odd we have:

$$G_{med}^\lambda = \frac{\lambda}{n\bar{y}} n_{med}^l(y_{med} - \bar{y}_{med}^l) + \frac{(1-\lambda)}{n\bar{y}} n_{med}^h(\bar{y}_{med}^h - y_{med})$$

Noting that $n_{med}^l = n_{med}^h = \frac{n-1}{2}$,

$$G_{med}^\lambda = \frac{n-1}{2n\bar{y}}[(1-\lambda)\bar{y}_{med}^h - \lambda\bar{y}_{med}^l]$$

If n is even there is no individual with median income since $y_1 < y_2 < \cdots < y_n$.

(iii) From (5) and (22), given that $n_n^h = 0$,

$$G_n = \frac{\lambda}{n\bar{y}}[n_n^l(y_n - \bar{y}_n^l)]$$

$$= \frac{1}{n\bar{y}}[(n-1)y_n - \sum_{j=1}^{n-1} y_j]$$

$$= \frac{\lambda}{n\bar{y}}[ny_n - \sum_{j=1}^{n} y_j]$$

$$= \frac{\lambda}{n\bar{y}}[ny_n - n\bar{y}]$$

$$= \lambda[\frac{y_n}{\bar{y}} - 1]$$