Working Paper No. 826

Tale of Two Ginis in the United States, 1921–2012

by

Markus P. A. Schneider*
University of Denver

Daniele Tavani†
Colorado State University

January 2015

* 2000 E. Asbury Ave., SH 246, Denver, CO 80208; markus.schneider@du.edu
† C306 Clark Building, Fort Collins, CO 80523; daniele.tavani@colostate.edu

Levy Economics Institute
P.O. Box 5000
Annandale-on-Hudson, NY 12504-5000
http://www.levyinstitute.org

ISSN 1547-366X

Abstract

Following a methodology proposed by Jantzen and Volpert (2012), we use IRS Adjusted Gross Income (AGI) data for the United States (1921–2012) to estimate two Gini-like indices representing inequality at the bottom and the top of the income distribution. We also calculate the overall Gini index as a function of the parameters underlying the two indices. Our findings can be summarized as follows. First, we find that the increase in the Gini index from the mid 1940s to the late 1970s seems to be mostly explained by an increase in inequality at the bottom of the income distribution, which more than offsets the decrease in inequality at the top. The implication is that middle incomes gained relative to high incomes, but especially relative to low incomes. Conversely, it is rising inequality at the top that appears to drive the rise in the Gini index since 1981. Second, inequality at the top of the income distribution follows a U-shaped trajectory over time, similar to the pattern of the share of top incomes documented by Piketty and Saez (2003, 2006) and Atkinson, Piketty, and Saez (2011). Third, the welfare effects of the different forces behind an increasing Gini index can be evaluated in light of the Lorenz-dominance criterion proposed by Atkinson (1970): both top-driven and bottom-driven increases in the index appear not to imply strict Lorenz dominance by previous income distributions, and therefore are not associated with lower welfare in an absolute sense. In a relative sense, however, once average growth rates over the two periods are taken into account, the top-driven increase in inequality since 1981 appears to have been welfare reducing.

Keywords: Income Distribution; Inequality; Gini Index

JEL Classifications: D3, D63

1 INTRODUCTION

Economic inequality, both regarding income and wealth, has surged to the top of academic and political debates in recent years. The contributions by Piketty and Saez (2003, 2006) are not only common references for economists, but often provide background information for opinion pieces in the popular press and for political arguments. Their work on compiling the World Top Income Database has allowed Atkinson et al. (2011) to track the patterns followed by top incomes worldwide. One of the most compelling features of these contributions lays in the documented U-shaped pattern followed by the share of top incomes in the US since the 1920s. Starting from the period between the two wars, when the top 1% were capturing up to 20% of total income, top income shares decreased to around 10% and remained stable at that level from the 1940s to the 1980s. A declining top income share during the early 20th century, and stability after WWII until the 1980s, would suggest declining inequality followed by a relatively stable level of inequality. Since the 1980s, the share of top incomes has increased steadily, till being back to pre-Great Depression levels, pointing to rising inequality.

Interestingly enough, it is hard to find support for such a U-shaped pattern when looking at the most common measure of inequality-the Gini index. The Gini has been steadily increasing, without much volatility, since the 1940s, thus pointing toward steadily increasing inequality in the United States.[1]

In this paper, we provide an analysis aimed at reconciling the contrasting evidence arising from the two measures of inequality. We utilize a methodology proposed by Jantzen and Volpert (2012), which exploits two interesting features of a well-fitting parametric Lorenz curve model: the apparent repeating patterns of how incomes are distributed among top incomes and bottom incomes. First, about half of all income goes to the top 10% of income earners, roughly half of that goes to the top 1%, half of that to the top 0.1%, and half of that goes to the top 0.01% (Atkinson et al., 2011), which implies that the Lorenz curve exhibits *right self-similarity* near the top of the distribution (Jantzen and Volpert, 2012). Second, Jantzen and Volpert (2012) further observe that the bottom 40% of the distribution appears to be approximately getting about a quarter of the income of the bottom 80%, and the bottom 20% is getting roughly a quarter of the bottom 40%, suggesting that the Lorenz curve is roughly *left self-similar* at the bottom of the distribution. They then propose a"‘hybrid model" that incorporates these behaviors asymptotically toward the respective ends of the distribution.[2] Further, Jantzen and

[1]Galbraith (2012) has provided a compelling illustration of the increase in the Theil index since the 1940s, as well as the comovements between the Theil and trends in the stock market in the US.

[2]The model Jantzen and Volpert (2012) proposed was actually already covered by Sarabia, Castillo, and Slottje (1999), who commented on its good fit of the data and derived conditions of Lorenz dominance.

Volpert (2012) exploit left-and right self-similarity to propose two Gini-like indices representing inequality at the top and at the bottom of the income distribution, respectively.

We expanded on the existing literature in estimating these two indices as well as the overall Gini using IRS Statistics of Income (SOI) data for estimated Adjusted Gross Income (AGI) over the period 1921-2012.[3] We are then able to evaluate the contribution of both indices to the increase in the US Gini since the 1940s, and find evidence for two different reasons for rising inequality in the US. The period 1940-1977 is marked by an increase in inequality at the bottom of the income distribution that trumps decreasing inequality at the top, while the period 1981-2012 is characterized by rising inequality at the top. Furthermore, our estimates of inequality at the top of the income distribution show a U-shaped pattern that closely mirrors the trajectory of the share of income going to the top 1% of earners in the US as documented by Piketty and Saez (2003, 2006).

It is important to provide intuition for the two measures of inequality corresponding to the two indices we estimate. On the one hand, the extent of right self-similarity, captured by our estimate for inequality at the top, implies a Pareto-distribution for the upper tail of the income distribution, which is a generally accepted feature.

On the other hand, interpreting inequality at the bottom as captured by left self-similarity deserves more exhaustive elaboration. In 1977, the bottom 80% of income earners captured 50% of total reported AGI. The middle two quintiles took home three quarters of the bottom 80%'s income share, leaving just under a quarter-about 11.6% of total AGI-for the bottom two quintiles. The second quintile also captured roughly three quarters of the total income going to the bottom 40%, leaving a quarter for the bottom quintile. Continuing onward, in 1977 the bottom decile received just over a quarter of the income share going to the bottom quintile. This represents the peak of inequality at the bottom in the period covered by our analysis. For contrast, in 1944 when inequality at the bottom by the Jantzen and Volpert (2012) measure was much less, the bottom two quintiles received approximately a third of the total income going to the bottom 80%, the bottom quintile received roughly a third going to the bottom 40%, and so on.

The implication of inequality at the bottom is that small moves upward in the distribution are associated with disproportionate gains in the share of income captured: moving from the bottom decile to the second decile implies a tremendous increase in the share of the nations' total income. The increase in inequality at the bottom we find from 1944 to 1977 means that the middle of the distribution was capturing a larger share of total income at the expense of the bottom percentiles. In other words, moving on up into the middle class represented substantial gains for those households that were able to do so,

[3]All estimates are reported in Table 3 in the Appendix.

while leaving less for those left behind.

Finally, the Lorenz curves corresponding to our estimates can be evaluated in light of the Lorenz-dominance criterion proposed by Atkinson (1970), according to which-ignoring welfare gains due to increasing real total income per capita-a strictly dominated Lorenz curve implies an absolute welfare loss in terms of income distribution independent of how"'inequality-averse" the social welfare function used to evaluate distributional outcomes is. Our results illustrate that both the period featuring an increase in inequality at the bottom and the more recent period with rising inequality at the top are characterized by crossing Lorenz curves, and therefore do not give rise to strict Lorenz dominance. The implication is that both periods of rising inequality are not necessarily characterized by welfare losses in an absolute sense. Yet it is possible to differentiate between the two periods in a *relative* sense, once we take into account the fact that the mean of the distribution changes over time. Comparing Lorenz curves for key years, we argue that the bottom-driven inequality increases between the 1940s and the end of the 1970s-a period characterized by strong real GDP growth averaging around 4.4% a year-are not welfare-reducing in a relative sense.[4] Conversely, the top-driven inequality increases between the 1980s and 2012, also coupled with a lower average growth rate of 2.8% per year, appear to be relatively welfare-reducing.

2 TWO GINIS

The method used to calculate both the overall level of inequality, and the low-and high-Gini coefficients, was proposed by Jantzen and Volpert (2012) based on the aforementioned left and right self-similarity of the observed Lorenz curve. On the one hand, right self-similarity is consistent with a Pareto-distribution of top incomes, and implies that the Lorenz curve has the form $L(x) = 1 - (1 - x)^q$, where x is the cumulative population proportion that receives the $L(x)$ cumulative share of income. On the other hand, left self-similarity implies the Lorenz curve given by $L(x) = x^p$. Combining left and right self-similarity, Jantzen and Volpert (2012) propose the following hybrid 2-parameter model of the Lorenz curve:

$$L(x; p, q) = x^p \ (1 - (1 - x)^q) \tag{1}$$

This model actually fits into the class of ordered parametric Lorenz curves

[4]Average growth, however, lowers to 3.1% per year if the sample is restricted to the 1945-1980 period.

discussed by Sarabia et al. (1999).[5] As it has been documented in both these contributions, the model performs well despite its apparent simplicity. The parameters p and q are easy to estimate from observed Lorenz curve coordinates using nonlinear least squares. Furthermore, an overall Gini coefficient based on (1) can be calculated directly from the estimated parameters, according to equation (2):

$$G = 1 - \frac{2}{p+1} + 2\,\frac{\Gamma(p+1)\Gamma(q+1)}{\Gamma(p+q+2)} \tag{2}$$

where $\Gamma(\cdot)$ denotes the Gamma function. For comparison, we also estimate the Gini directly from the data used for this study using a typical trapezoidal approximation to the area under the Lorenz curve.

Jantzen and Volpert (2012) show that this model allows for the direct calculation of an inequality index G_0 based solely on the level and extent of left self-similarity, and an inequality index G_1 based on right self-similarity.[6] In this sense, G_0 represents inequality at the bottom of the observed income distribution and G_1 captures inequality at the top by directly reflecting the fattening of the power-law tail.

$$G_0 = \frac{3p}{p+2} \tag{3}$$

$$G_1 = \frac{1-q}{1+q} \tag{4}$$

Note that G_0 is a strictly increasing function of p, while G_1 is strictly decreasing in q.

2.1 Interpretation

The coefficients given in (3) and (4) are Gini-like indices of inequality in that a larger number indicates greater inequality and that they are bounded between zero and one. The parametric Lorenz curve that gives rise to both, given in (1), implies that the distribution asymptotically follows a simple power distribution near $x = 0$ and a classical Pareto distribution at the upper end (near $x = 1$).[7]

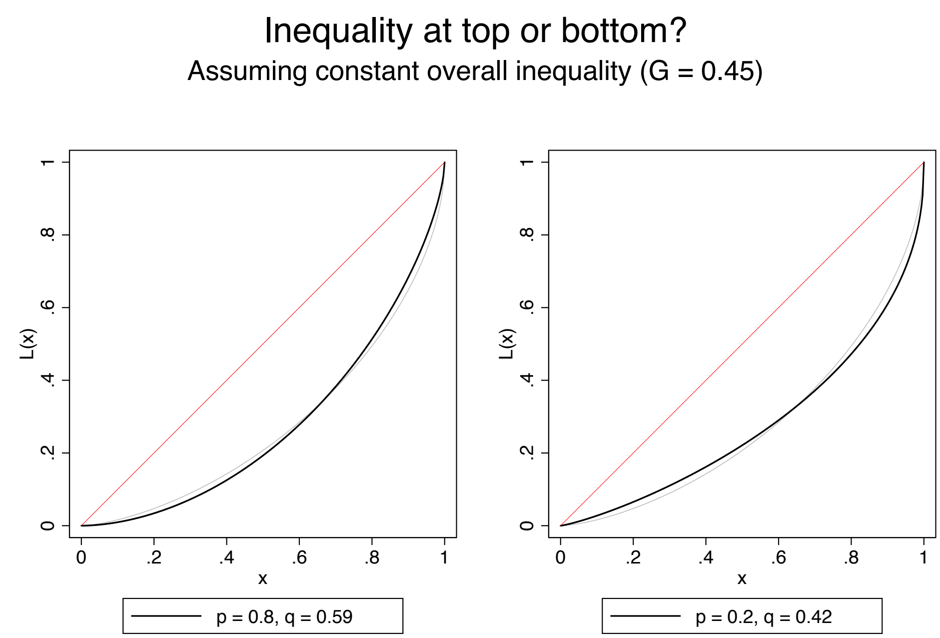
[5] However, the earlier authors did not consider the asymptotic features with respect to self-similarity, nor did they develop the Gini-like indices for inequality at the top and bottom. Also notable is that fitting a parametric Lorenz curve is, in a sense, more general than fitting a specific parametric distribution as has been recent practice in some papers, as Sarabia et al. (1999) point out.

[6] Our formulation of G_0 differs from that in Jantzen and Volpert (2012) by a factor of three to ensure that $G_0 \in [0, 1]$.

[7] Asymptotic right self-similarity is also consistent with recently popular parametric distribution such as the GB2 or Dagum (see Kleiber and Kotz, 2003).

As suggested in the Introduction, the interpretation of a high level of inequality at the bottom of the distribution requires some clarification. Basically, a high G_0 (closer to 1) implies that the very bottom percentiles receive a smaller share of total income. Specifically, the bottom decile receives only a fifth of the share of income going to the bottom quintile, and the bottom 5% receive only a fifth of that, and so on. A large G_0 implies that the poor are relatively poorer compared to the middle, but at the same time there are fewer individuals making very low incomes in absolute terms. Figure 1 and Table 1 illustrate the difference between a distribution characterized by high levels of inequality at the bottom versus high inequality at the top. When G_0 is close to zero, then the Lorenz curve near $x = 0$ is close to linear implying that the income share going to the bottom x% of the population rises roughly proportionally with x, which can be clearly seen in the right panel of Figure 1.

Figure 1: Inequality at the Bottom versus at the Top



Note: Comparing the Lorenz curve for $p = q = 0.5$ (shown in light gray) to those when inequality is high at the bottom (left) or top of the distribution (right), while overall inequality is held constant.

It is only a minor oversimplification to think of high inequality at the bottom of the distribution (high G_0) and low inequality at the top as characteristics of a strong middle class: successful attempts to pull oneself out of poverty (or social policies aimed at aiding individuals out of poverty) pay off in spades. Gains beyond that, however, are both modest in scale and frequency. By contrast, a low G_0 but high G_1 (right panel in Figure 1)

Table 1: Income Shares and Inequality

	Share of Income going to ...			
	... bottom 20%	**... middle 60%**	**... top 20%**	**... top 1%**
Baseline ($p = q = 0.5$)	4.7%	44.7%	50.6%	10.5%
High G_0 ($p = 0.8$)	3.4%	47.9%	48.7%	7.3%
High G_1 ($q = 0.2$)	6.5%	40.7%	52.8%	14.4%

Note: Holding the overall level of inequality as measured by the Gini coefficient constant, but varying whether inequality is more pronounced at the top or bottom of the distribution, implies stark differences in the share of income going to the bottom, middle, and top quintiles-and the top 1%.

characterizes a distribution where most individuals share relatively little income quite equitably, while a lot of income is reserved to reward top earners exceptionally well.

2.2 Lorenz Dominance and Welfare

Sarabia et al. (1999) show that there is a straightforward Lorenz ranking that can be established for the model given in (1). We say that distribution X Lorenz-dominates Y, formally $X \geq_L Y$, if and only if $L_X(u) \geq L_Y(u) \, \forall \, u \in [0, 1]$ so that the distribution of X is less unequal and welfare superior under the conditions spelled out in Atkinson (1970).[8] Hence, if two distributions give rise to crossing Lorenz curves, they cannot be ranked in terms of welfare without reference to a specific social welfare function and the associated level of inequality aversion.

Two distributions can be ranked in terms of the parameters p and q of their respective Lorenz curves when the following conditions hold:

$$L(x; p_1, q_1) \geq L(x; p_2, q_2) \text{ if and only if } p_2 > p_1, q_2 \leq q_1 \text{ or } p_2 \geq p_1, q_2 < q_1$$

However, the criteria of welfare loss due to inequality as it relates to Lorenz dominance spelled out by Atkinson (1970) presumes that the distributions being compared have the same mean. This is clearly not the case when comparing income distributions for over nine decades. Instead, we consider only the *relative* welfare loss due to changing inequality and leave aside the possibly off-setting welfare gains due to rising real per capita income.

[8] Specifically, we need only assume a smooth social welfare function that exhibits diminishing returns to income.

3 DATA

We use the IRS Statistics of Income (SOI) data for estimated Adjusted Gross Income (AGI) by size of AGI published as Table 1.1 in the annual IRS SOI reports (see Internal Revenue Service, 1995, for example). AGI is a widely used measure of pre-tax income from all sources. This dataset has two advantages over the Census Bureau data used by Jantzen and Volpert (2012), the first being that the series dates all the way back to 1921. Even taking into account some procedural changes to the tax code and reporting of SOI data, a consistent series can easily be constructed from 1944 to 2012. A second advantage is that the data covers the upper tail of the income distribution much better than the top-coded income reports available from the Census. Finally, the fact that the SOI data is binned gives it a very usable format from which it is trivial to construct the Lorenz curve coordinates necessary to estimate p and q.

One important point to note is that the IRS SOI only captures individuals who filed tax returns regardless of whether they reported any taxable income. Our estimates of inequality in pre-tax income specifically only covers returns that reported \$1 or more of income. That means that our measure of inequality neglects those who earned nothing or reported losses, and therefore understates the more general social level of inequality in some sense.

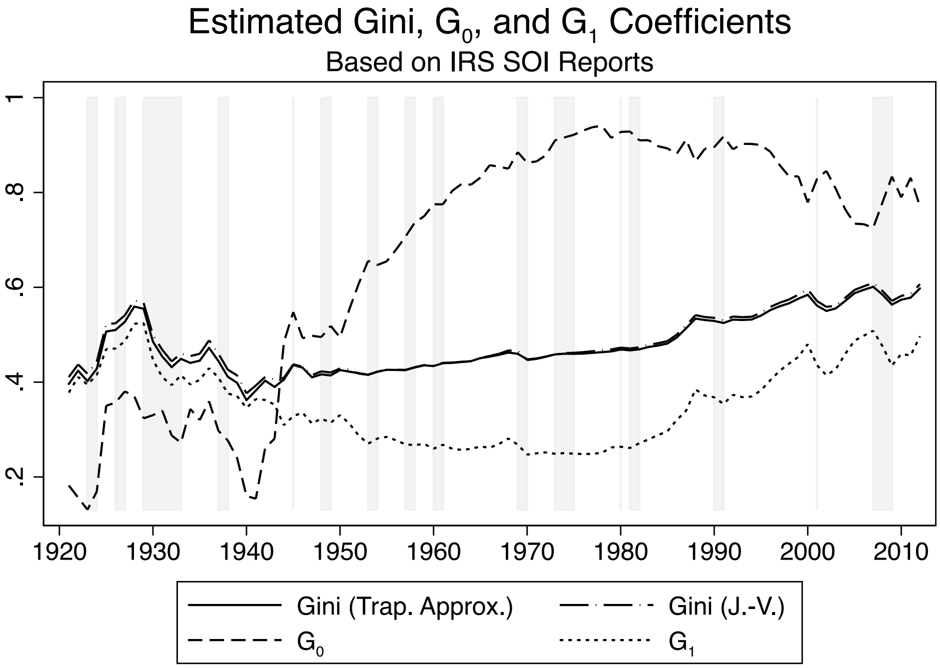
The two parameters p and q in (1) were estimated using Stata's nonlinear least squares command, `nl`. Consistent with the findings by Sarabia et al. (1999) and Jantzen and Volpert (2012), the fit was very good across years with reported values of $R^2 \geq 0.9995$.

4 RESULTS

Figure 2 shows the estimated series of the overall Gini calculated using (2), the G_0 and G_1 indices, and an overall Gini estimate based on a trapezoidal approximation for the area under the Lorenz curve. It is clear that the overall Gini estimated from the parameters of the parametric Lorenz curve is consistent with the trapezoidal approximation. While overall inequality measured by the Gini shows a persistent upward trajectory from the mid 1940s to the present, the movements in G_0 and G_1 reveal that this trend is not driven by the same underlying changes in the distribution during that period.

One of the notable discrepancies in the recent literature is that the trend in inequality measured using the Gini does not seem to reflect the patterns in income distribution uncovered by Piketty and Saez's work: specifically, the rise in the share of income going to the top 1% income earners rose much faster recently than inequality measured by the Gini, and the Gini does not show the same U-shaped trajectory during

Figure 2: Two New Ginis

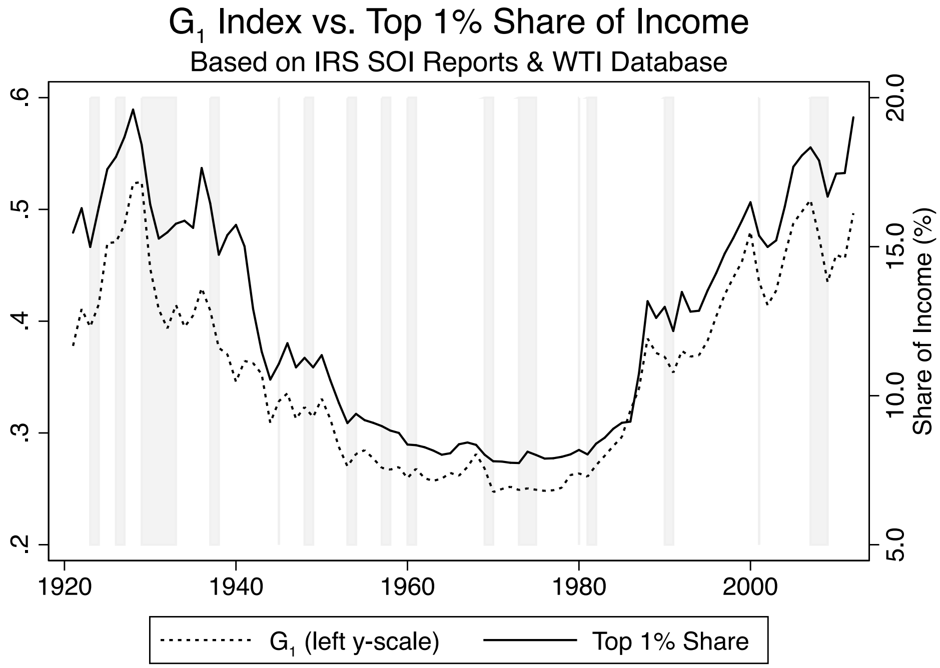


Note: Two estimates of the overall Gini coefficient, and G_0 and G_1 are shown in the figure. Clearly, the overall Gini estimated using (2) is very similar to the Gini estimated based on the trapezoidal approximation of the area under the Lorenz curve.

the 20th century one finds in Piketty and Saez (2003, 2006). The reason for this discrepancy is that the overall Gini hides the fact that during different decades in the 20th century, changes in different parts of the distribution of income were responsible for driving the rise in overall inequality. Figure 3 plots the isolated series of G_1 together with the share of income going to the top 1%, and shows a good qualitative correspondence between the two. Tracking changes in the Lorenz curve that imply a bowing out at the top (like that illustrated in the right panel of figure 1) separately from changes at the bottom reveals that those changes closely correspond to the changes in top income shares documented by Piketty and Saez (2003, 2006); Atkinson et al. (2011).

By contrast, rising inequality was driven by changes at the bottom of the income distribution from roughly WWII until the end of the 1970s. As discussed earlier and is well known, the post-war period is characterized by distributional gains at the middle. In terms of our analysis, those gains came primarily at the expense of those who remained at the bottom, though the modest decline in G_1 indicates that they also came at the expense of those at the top of the distribution.

Figure 3: Evolution of G_1 versus the Income Share of the Top 1%



4.1 Regime Changes

A simple time series analysis suggests that, while the overall Gini appears to be increasing along a fairly constant trend, the changes in inequality over the period 1944-2012 can be separated into two regimes. There appear to be two breaks in the series of the parameter q-1977 and 1981-and one in the series of parameter p-1977.[9] Prior to 1977, q was on average increasing.[10] Table 2 show the estimated average first differences for each sub-period.

Table 2: Average First Differences

Period	Δq	Δp
1944-1977	0.004	**0.021**
1977-1981	-0.004	-0.006
1981-2012	**-0.008**	

Note: Average first differences before and after the most likely breaks in the estimates series. Statistically significant first differences appear in bold.

[9] Based on sequential Chow tests for breaks and choosing the year(s) with the largest test statistic (i.e., an informal Quandt Likelihood Ratio test). Break dates are based on $\Delta q_t = \beta_0 + u_t$ and $\Delta p_t = \gamma_0 + v_t$. Formal tests suggested that first differences of the parameters were stationary and no further lags were necessary.

[10] Large fluctuations mean that its first difference is not statistically different from zero.

The on-average increase of q from 1944 until 1977 implies a slow decline in inequality at the top as measured by G_1, which is apparent in Figure 2. After 1977, the trend in q reversed and the average first difference was negative (and statistically different from zero at the 2% level), consistent with the observation that inequality at the top was rising. A cautious appraisal suggests that inequality at the top of the income distribution was relatively stable from the end of WWII until 1981, but has been increasing ever since.

By contrast, p was increasing at a statistically significant (at the 1% level) rate prior to 1977, but effectively stabilized (and perhaps began a slow decline) thereafter. From the end of WWII until the late 70s, increasing inequality as measured by the Gini was driven by inequality at the bottom as the distance between low-and mid-level incomes grew.

Considering these changes in light of the ordering suggested by Sarabia et al. (1999) and the welfare implications spelled out by Atkinson (1970), the impact due to changes in inequality from 1944 to 1977 is ambiguous and depends on a society's degree of inequality aversion. Restricting the attention to distributions at the end-points of that period, the Lorenz curve for the bottom 77% of distribution in 1977 lies strictly below that of that of the 1945 Lorenz curve, while the top 23% of income earners actually experienced a reduction in inequality, as can be seen in Figure 4a. In other words, the bottom 77% of income earners shared 46% of total income less equitably in 1977 than in 1945, because a larger portion of it was concentrated near the middle of the distribution.

At first pass, it seems that the story for the period from 1981 until 2012 is similar: decreases in both q and p suggest again that consecutive Lorenz curves do cross, and that therefore one's judgment about the welfare implications depends on the degree of inequality aversion. When we consider where the Lorenz curves for 1981 and 2012 cross, however, the welfare conclusions seem less ambiguous. For only the bottom 6.6% of income earners was the 1981 Lorenz curve below that of the 2012 curve, suggesting that only the bottom 0.3% of income was more equitably distributed in 2012 than in 1981. The vast majority of the 2012 Lorenz curve lies below that of 1981 and for over 93% of income earners inequality had unambiguously increased (see Figure 4b). An illustrative comparison is that by 2012, the bottom 77% of income earners shared only about 33% of total income. Only gross indifference toward inequality would suggest the latter represents a welfare improvement due to distributional changes.

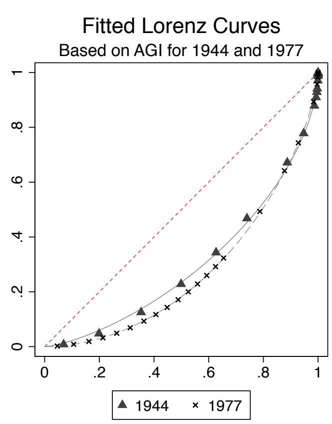
In fact, we find that the 2012 Lorenz curve is Lorenz-dominated (i.e., lies strictly below the Lorenz curves for other years) by all distributions prior to 1960 except 1928 and 1929, suggesting that, in Atkinson's sense, the current level of inequality reflects a distribution of income associated with a strictly inferior social welfare compared to almost any year prior to 1960, when welfare gains due to increases in real mean income are not taken into account. Attesting to this is the notably further movement of the Lorenz curve in the last 31 years of the sample compared to the first 37 as seen in Figure 4.

The referenced welfare changes are, of course, relative due only to changes in the distribution of income and ignoring increases in total income. However, it is not clear that there are total income gains of sufficient magnitude in the latter period to offset the relative negative impact described. As mentioned in the Introduction, real GDP growth from WWII until the 1970s was somewhat higher (4.4%) than from the 1980s until the present (2.8%). While the average growth rate is quite sensitive to the choice of start and end years, there is little evidence that the latter period was characterized by broadly shared gains from growth that would justify putting up with the changes in distribution highlighted here and documented by Atkinson et al. (2011).

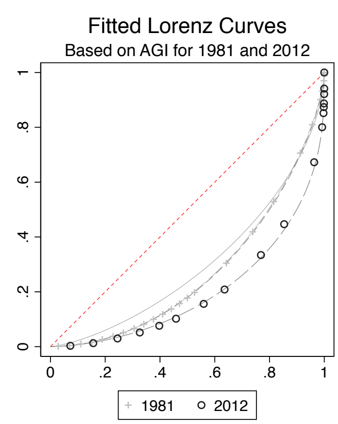
These results offer a far more nuanced picture of how the distribution of income has changed than could be gleaned from changes in the Gini alone. The Gini series suggests near-constant trend of increasing inequality from WWII to the present. Breaking down the Gini into two indices has enabled us to discriminate between an increase in overall inequality driven by the growing distance of middle-class incomes from the bottom, from an increase in overall inequality driven by the growing share of income concentrated at the very top. For this reason, our results complement those documented by Atkinson et al. (2011) and earlier work on the share of top incomes.

Figure 4: Fitted Lorenz Curves

(a) 1944 and 1970



(b) 1981 and 2012



Note: The Lorenz curves for 1944 and 1977 are shown in gray and without data markers in (b).

5 CONCLUSION

This paper has tackled the discrepancy between the U-shaped pattern followed in post-war United States by the share of top incomes on the one hand, and the steadily increasing Gini index on the other. We used the methodology suggested by Jantzen and Volpert (2012) to evaluate the trend in the Gini in light of Gini-like indices representing inequality at the bottom and at the top of the income distribution. In doing so, we made use of the SOI data for AGI, which has enabled us to work with a considerably longer sample period relative to previous contributions.

We found that bottom-driven inequality appears to explain the rising Gini up to 1977, thus signaling gains in middle incomes both at the expenses of top incomes *and especially bottom incomes*; while it is almost entirely gains at the top at the expense of both the middle and much of the bottom that drives the increase in the Gini after 1981.

Moreover, we documented that the time pattern followed by our estimates of inequality at the top closely mirrors the U-shaped trajectory of the share of top incomes found in Piketty and Saez (2003, 2006); Atkinson et al. (2011).

Finally, we provided an evaluation of subsequent distributions of income (1944, 1977, 1981, and 2012) in light of the Lorenz-dominance criterion, and found that the most recent distribution of income is almost entirely dominated by the distribution at the beginning of the 1980s, and strictly dominated by almost any distribution before 1960. These back-of-the-envelope welfare considerations, however, should be considered preliminary.

References

Atkinson, A. B. (1970). On the measurement of Inequality. *Journal of Economic Theory 2*(3), 244–263.

Atkinson, A. B., T. Piketty, and E. Saez (2011). Top Incomes in the Long Run of History. *Journal of Economic Literature 49*(1), 3–71.

Galbraith, J. K. (2012). *Inequality and Instability*. Oxford University Press.

Internal Revenue Service. Statistics of Income (SOI) Reports (1996 - 2012). Accessed August 2014: `http://www.irs.gov/uac/SOI-Tax-Stats-Individual-Income-Tax-Returns-Publication-1304-%28Complete-Report%29`.

Internal Revenue Service. Statistics of Income (SOI) Reports Archive (1934 - 1953). Accessed August 2014: `http://www.irs.gov/uac/SOI-Tax-Stats-Archive---1934-to-1953-Statistics-of-Income-Report,-Part-1`.

Internal Revenue Service. Statistics of Income (SOI) Reports Archive (1954 - 1999). Accessed August 2014: `http://www.irs.gov/uac/SOI-Tax-Stats-Archive-1954-to-1999-Individual-Income-Tax-Return-Reports`.

Internal Revenue Service (1995). Statistics of Income - 1995: Individual Tax Returns. Technical Report 1304, Department of the Treasury.

Jantzen, R. T. and K. Volpert (2012). On the mathematics of income inequality: Splitting the gini index in two. *The American Mathematical Monthly 119*(10), 824–837.

Kleiber, C. and S. Kotz (2003). *Statistical Size Distributions in Economics and the Actuarial Sciences*. John Wiley.

Piketty, T. and E. Saez (2003). Income Inequality in the United States, 1913 - 1998. *Quarterly Journal of Economics 118*(1), 1–39.

Piketty, T. and E. Saez (2006). The Evolution of Top Incomes: A Historical and International Perspective. *American Economic Review 96*(2), 200–205.

Sarabia, J.-M., E. Castillo, and D. J. Slottje (1999). An ordered family of Lorenz curves. *Journal of Econometrics 91*(1), 43–60.

APPENDIX

Table 3: Parameter and Index Estimates, 1921-2012

Year	p	q	Gini	G_0	G_1	Year	p	q	Gini	G_0	G_1
1921	0.129	0.451	0.411	0.182	0.378	1967	0.796	0.575	0.460	0.854	0.270
1922	0.110	0.418	0.437	0.157	0.411	1968	0.791	0.561	0.467	0.850	0.281
1923	0.091	0.434	0.418	0.131	0.395	1969	0.837	0.577	0.466	0.885	0.268
1924	0.120	0.413	0.443	0.169	0.415	1970	0.807	0.604	0.447	0.863	0.247
1925	0.264	0.361	0.522	0.350	0.470	1971	0.811	0.600	0.449	0.865	0.250
1926	0.270	0.360	0.524	0.357	0.471	1972	0.828	0.598	0.453	0.878	0.252
1927	0.290	0.346	0.541	0.380	0.486	1973	0.871	0.601	0.459	0.910	0.249
1928	0.282	0.313	0.572	0.371	0.523	1974	0.879	0.599	0.461	0.916	0.250
1929	0.242	0.312	0.567	0.324	0.525	1975	0.887	0.601	0.462	0.922	0.249
1930	0.248	0.382	0.499	0.330	0.447	1976	0.899	0.602	0.463	0.931	0.248
1931	0.256	0.418	0.468	0.340	0.410	1977	0.909	0.602	0.465	0.938	0.249
1932	0.212	0.435	0.444	0.287	0.393	1978	0.913	0.599	0.467	0.941	0.251
1933	0.200	0.414	0.460	0.272	0.415	1979	0.878	0.584	0.469	0.915	0.262
1934	0.259	0.434	0.455	0.344	0.395	1980	0.895	0.583	0.473	0.928	0.264
1935	0.239	0.423	0.460	0.320	0.405	1981	0.896	0.586	0.471	0.929	0.261
1936	0.272	0.399	0.488	0.360	0.430	1982	0.871	0.573	0.473	0.910	0.271
1937	0.220	0.419	0.460	0.297	0.410	1983	0.871	0.563	0.479	0.910	0.280
1938	0.204	0.453	0.426	0.277	0.376	1984	0.854	0.552	0.482	0.898	0.288
1939	0.174	0.460	0.414	0.240	0.370	1985	0.848	0.543	0.486	0.893	0.296
1940	0.113	0.486	0.376	0.160	0.346	1986	0.832	0.515	0.500	0.882	0.320
1941	0.108	0.466	0.393	0.154	0.364	1987	0.873	0.493	0.519	0.911	0.340
1942	0.191	0.468	0.411	0.262	0.362	1988	0.811	0.445	0.541	0.865	0.385
1943	0.206	0.479	0.406	0.280	0.353	1989	0.846	0.458	0.537	0.892	0.372
1944	0.385	0.528	0.409	0.484	0.309	1990	0.851	0.462	0.536	0.895	0.368
1945	0.447	0.506	0.438	0.548	0.328	1991	0.882	0.477	0.530	0.918	0.354
1946	0.393	0.498	0.432	0.493	0.335	1992	0.845	0.456	0.538	0.891	0.373
1947	0.398	0.524	0.415	0.498	0.313	1993	0.861	0.462	0.537	0.903	0.368
1948	0.395	0.512	0.423	0.495	0.323	1994	0.860	0.460	0.538	0.902	0.370
1949	0.418	0.522	0.420	0.519	0.314	1995	0.857	0.447	0.546	0.900	0.383
1950	0.395	0.503	0.429	0.494	0.331	1996	0.838	0.424	0.559	0.886	0.404
1951	0.453	0.524	0.426	0.554	0.312	1997	0.801	0.405	0.567	0.858	0.424
1952	0.508	0.553	0.418	0.608	0.287	1998	0.770	0.390	0.574	0.834	0.439
1953	0.560	0.575	0.415	0.656	0.270	1999	0.770	0.376	0.584	0.834	0.453
1954	0.550	0.561	0.422	0.647	0.281	2000	0.701	0.351	0.596	0.779	0.480
1955	0.558	0.557	0.426	0.655	0.285	2001	0.764	0.393	0.571	0.829	0.435
1956	0.586	0.566	0.426	0.680	0.277	2002	0.785	0.414	0.559	0.846	0.414
1957	0.617	0.576	0.426	0.707	0.269	2003	0.737	0.401	0.561	0.808	0.427
1958	0.651	0.578	0.431	0.737	0.267	2004	0.684	0.371	0.578	0.765	0.459
1959	0.667	0.575	0.436	0.750	0.269	2005	0.648	0.344	0.594	0.734	0.488
1960	0.697	0.588	0.435	0.775	0.260	2006	0.647	0.335	0.602	0.733	0.498
1961	0.697	0.577	0.441	0.775	0.268	2007	0.638	0.326	0.609	0.725	0.508
1962	0.732	0.588	0.441	0.804	0.259	2008	0.702	0.355	0.592	0.780	0.476
1963	0.750	0.591	0.443	0.818	0.257	2009	0.769	0.394	0.571	0.834	0.435
1964	0.748	0.588	0.444	0.816	0.259	2010	0.714	0.370	0.582	0.789	0.459
1965	0.767	0.582	0.451	0.831	0.264	2011	0.766	0.374	0.586	0.831	0.456
1966	0.801	0.585	0.455	0.858	0.262	2012	0.691	0.336	0.607	0.770	0.497