A Note on New Estimates of the Distribution of Income in the 1920s

GENE SMILEY

The period from 1916 through 1929 is important to anyone interested in the history of income inequality in the United States. Current estimates suggest that income inequality increased in the antebellum period and reached a “plateau of high income inequality from the Civil War to 1929.” Inequality seems to have risen from the turn of the century to 1916 and then declined sharply during the First World War. The 1920s brought a new surge of income inequality which, by 1929, pushed inequality to about the 1916 level. From 1929 to 1950 there was a significant decrease in income inequality in the United States followed by a period of relative distributional stability. Thus, 1916 and 1929 appear to have marked peaks in income inequality in the history of the United States.

Because it directly preceded the Great Depression the surge of income inequality during the prosperous 1920s has generated considerable interest. In 1946 Arthur Burns developed data which provided preliminary measures of this growing disparity, but the most important evidence on changes in the distribution of income during the 1920s was presented in 1953 in Simon Kuznets’s pathbreaking study, Shares of Upper Income Groups in Income and Savings. Kuznets carefully constructed income shares for upper-income percentiles of the total and nonfarm population. These shares provided evidence of increasing income inequality. For example, the share of disposable income of the top 1 percent of the nonfarm population rose from 13.13 percent in 1919 to 19.07 percent in 1929. From the business-cycle peak in 1923 to the cyclical peak in 1929 this share rose from 12.78 to 19.07 percent. Similar, though less pronounced, changes were found for the second and third, fourth and fifth, and sixth and seventh percentiles. The share of disposable income for the lower 93 percent of the nonfarm population fell from 71.00 percent in 1919 to 69.74 percent in 1923 to 61.29 percent in 1929.

In 1977 Charles Holt used Kuznets's data to estimate real per capita disposable incomes for nonfarm income percentiles. Holt found that the real per capita disposable incomes of the lower 93 percent of the nonfarm population fell 4 percent and concluded that, “The inference is clear. Consumption increased markedly during the New Era in the face of little or no improvement in the consumption habits of the bulk of the nonfarm population.”

Recent research has shown that Kuznets’s share estimates, and Holt’s real per capita income estimates, overstate the rise in inequality during the twenties because no adjustments were made for changes in tax-avoidance behavior. This article draws upon new
estimates of income shares in the 1920s to re-examine income distribution in the United States over that decade.

THE NEW ESTIMATES

Simon Kuznets developed his income-share estimates from data reported in federal income-tax returns and from separate estimates of aggregate income. The Statistics of Income reported the aggregate income for various net-income tax classes in addition to the population represented by the returns in each net-income tax class. Kuznets calculated the per capita incomes for each class and arrayed these from the highest to the lowest net income per capita. For each class he also calculated the share of total population and share of total income it represented. These were then cumulated from the highest per capita income down and interpolated at 1, 3, 5, and 7 percent of the nonfarm population. This yielded the share of total income received by the top 1 percent, second and third percentiles, fourth and fifth percentiles, and sixth and seventh percentiles of the nonfarm population. The share of the lower 93 percent of the nonfarm population was calculated as the residual. Kuznets made adjustments to correct for the compensation of nonfederal government employees, to include imputed rent for the upper-income percentiles, to exclude federal income taxes paid, to include the excess of realized gains over realized losses from the sales of assets, to correct for variations in family status among the upper-income taxpayers, and to correct for unwarranted inclusions and unwarranted deductions in income reported on the tax returns.6

The problem is that over the course of the twenties some of the additional income reported by the upper-income recipients on their tax returns came not from additional income they received, but from income that previously was not taxable and therefore was not reported.7 Obtaining more accurate estimates of income shares required eliminating the changes in reported income for the upper-income percentiles that arose from changes in their income-tax-sheltering behavior. The method by which this was accomplished can be summarized here.8 Due to income-tax avoidance the Statistics of Income reported too few returns in the higher-income classes. This was especially true in the early twenties when marginal income-tax rates were much higher. As marginal income-tax rates declined and wealth was shifted from tax-avoiding assets and activities into taxable assets and activities, the reported number of returns and total income in all upper-income classes rose. Kuznets assumed that all of this growth represented additional income received by the upper-income classes. However, we know that some of this was not additional income but income that these taxpayers had previously been receiving but had not reported because it was exempt from federal personal income taxes.9

The income shares for the lower 93 percent of the nonfarm population could not be estimated directly. Rather Kuznets estimated these as a residual by subtracting the estimated share of the top 7 percent of the nonfarm population from 100. If we assume that Kuznets’s estimates of aggregate income are correct then this procedure underestimated the

---

6 Nonfederal government employees were exempt from federal income taxes.
8 This summary is based upon Smiley, “New Estimates,” and “Technical Appendix,” pp. 2–14, where it is described in more detail and which contains the supporting tables as well as tables reporting the new estimates.
9 We do not know how much of the avoidance was legal and how much was illegal. The discussion in the 1917 to 1925 period emphasized legal tax avoidance. Our adjustment method does not, in fact, rely on whether the avoidance was legal or illegal.
share of income received by the upper income percentiles in the high-marginal-tax-rate years at the beginning of the twenties. This also means it overestimated the share of income received by the bulk of the nonfarm population in the early twenties. As marginal income-tax rates fell during the twenties these over- and underestimates were reduced. Therefore, the rise in the shares of the upper-income percentiles during the twenties was overstated and the fall in the share of the bulk of nonfarm population during the twenties was also overstated.

To correct this, the number of returns, income, and population that would have been reported in each income-tax bracket in the absence of changes in income-tax-avoidance behavior were required. In other words these are the number of returns, income, and population which would have been reported if marginal tax rates in each income-tax class had been constant over the entire 1919 through 1929 period. The first step in the process was to estimate how the number of reported returns varied in each net-income tax class as marginal tax rates varied while controlling for other relevant variables. With this estimated relationship for each tax class, the marginal tax rate could be held constant while the other variables took their actual values so as to estimate what the number of returns in each tax class would have been with constant marginal tax rates.

The marginal federal personal income-tax rates of 1916 were used for this. Income components for each net-income tax class for each year were estimated by multiplying the Statistics of Income data by the ratio of the new estimates of the number of returns to the reported number of returns. Adding together these new estimates of income components provided revised estimates of total income for each net-income tax class for each year. Using Kuznets’s data the population per return for each net-income tax class for each year could be estimated and this provided estimates of the population represented by the adjusted returns in each net-income tax class for each year. These shares were then adjusted for the compensation of nonfederal government employees, for imputed rent, for federal income taxes, for family status, for unwarranted inclusions, and for the maximum effect of unwarranted inclusions and deductions following the same procedures Kuznets devised, but using the new data. Subtracting the percentage shares of the top seven percentiles from 100 provided the new estimates of the shares of disposable income for the lower 93 percent of the nonfarm population for each of these years.

The new estimates of income shares generated by this process showed some significant differences from Kuznets’s estimates. The income shares for the top percentiles grew much less rapidly between 1923 and 1929. In particular the income share of the top percentile grew only 58 percent as rapidly as what Kuznets found. The income share for the lower 93 percent fell less than half of what Kuznets found and the real per capita incomes of the lower 93 percent of the nonfarm population rose rather than fell between 1923 and 1929.

It is easier and more accurate to observe the overall changes in income distribution by constructing Gini Coefficients. Table 1 presents the Gini coefficients for Kuznets’s and for the revised share estimates from 1919 through 1929. As can be seen the estimates revised

10 The other variables that were included in the regressions were real GNP per capita, the consumer price index, and the size of the labor force. The marginal tax rate was lagged one year in the regressions. For a detailed discussion of this see Smiley, “Technical Appendix,” pp. 2–14.

11 The income share for the top 1 percent rose 3.671 percent between 1923 and 1929 while Kuznets found it to rise by 6.29 percent. The income share of the lower 93 percent fell 3.979 percent between 1923 and 1929 while Kuznets found it to fall 8.45 percent. The real per capita incomes of the top 1 percent increased 29.17 percent while Kuznets’s data showed their real per capita incomes to rise 63.27 percent between 1923 and 1929. The real per capita incomes of the lower 93 percent of the nonfarm population increased 2.43 percent (or by about $3.21 per year) between 1923 and 1929 rather than falling 3.89 percent as Kuznets’s data showed.
New Estimates of Income Distribution

TABLE 1
GINI COEFFICIENTS, 1919–1929

<table>
<thead>
<tr>
<th>Year</th>
<th>Kuznets</th>
<th>Revised Excluding Capital Gains</th>
</tr>
</thead>
<tbody>
<tr>
<td>1919</td>
<td>0.2610</td>
<td>0.3590</td>
</tr>
<tr>
<td>1920</td>
<td>0.2200</td>
<td>0.3350</td>
</tr>
<tr>
<td>1921</td>
<td>0.3152</td>
<td>0.4382</td>
</tr>
<tr>
<td>1922</td>
<td>0.3319</td>
<td>0.4198</td>
</tr>
<tr>
<td>1923</td>
<td>0.2965</td>
<td>0.4451</td>
</tr>
<tr>
<td>1924</td>
<td>0.3188</td>
<td>0.4219</td>
</tr>
<tr>
<td>1925</td>
<td>0.3566</td>
<td>0.4291</td>
</tr>
<tr>
<td>1926</td>
<td>0.3463</td>
<td>0.4534</td>
</tr>
<tr>
<td>1927</td>
<td>0.3653</td>
<td>0.4407</td>
</tr>
<tr>
<td>1928</td>
<td>0.3990</td>
<td>0.4324</td>
</tr>
<tr>
<td>1929</td>
<td>0.3906</td>
<td>0.4828</td>
</tr>
</tbody>
</table>

Trend Coefficient: 1919 to 1929 0.0146* 0.0097* 0.0011
Trend Coefficient: 1923 to 1929 0.0161* 0.0052 -0.0041
Percentage Change: 1923 to 1929 31.74 8.47 -4.23

* = Significantly different from zero at the 1 percent significance level. (The significance test is based on a two-tail t test.)

Note: The trend coefficient is $\beta$ in the regression, $GINI = \alpha + \beta(YEAR)$, where $GINI$ is the Gini Coefficient and $YEAR$ is the year for that percentage income share.


to eliminate changes in tax avoidance behavior show a greater amount of inequality for the nonfarm population in each year in the decade. However, the rise in inequality, which is the focus of this analysis, is clearly smaller with the revised estimates. Between 1919 and 1929 the Gini coefficient for the revised estimates rose 0.0097 per year while the Gini coefficient for Kuznets’s estimates rose 0.0146 per year. Between 1923 and 1929 the coefficient rose by 0.0161 per year for Kuznets’s estimates but only 0.0052 per year for the revised estimates. Clearly, adjusting income share estimates to eliminate changes due to changes in tax-avoidance behavior results in a considerably slower rise in income inequality during the decade of the 1920s.

SPECULATIONS ON THE SOURCES OF INEQUALITY CHANGES

These new estimates suggest that the increase in income inequality during the twenties was considerably less than Kuznets’s unadjusted income shares suggested. An obvious question is why this rise in inequality occurred. Two likely candidates for such a trend are easily found. First, there is evidence that during the First World War various aspects of the federal government’s controls tended to reduce incomes received by the upper-income groups. This would temporarily have created greater income equality during the 1917 to 1919 period. Second, the upper-income percentiles received much more of their income from investments and much less from wages. This was especially true for the highest income percentile. It would be expected that the securities-market boom of the last half of the twenties would have led to larger increases in the incomes of the upper-income percentiles.
The First World War

Jeffrey Williamson and Peter Lindert have argued that the First World War interrupted a long-term trend of rising inequality. Their analysis was based in part on the relationship between measures of income distribution and wage differentials between skilled and unskilled labor. This analysis complemented Lee Soltow’s examination of inequality among the upper-income groups. Soltow used income-tax-return data that limited the analysis to 1866 to 1871, 1894, and 1913 on. Though Soltow found a slight decline in upper income inequality from the 1866 to 1871 period to the 1913 to 1916 period, he also found a sharp decline in upper income inequality by 1918 and 1919, a decline that continued into the early twenties. Upper income inequality rose sharply in the 1925 to 1929 period.

As Williamson has pointed out, this sharp decline in inequality during the First World War “coincided with America’s first experiment with price controls.” In 1920 P. W. Garrett pointed out that Congress had set up machinery to control food and fuel prices to protect labor and the middle class from “exorbitant” increases in the prices of staple commodities. Government intervention actually began in 1916 when the Adamson Act mandated a reduction in the work-day in the railroad industry with no reduction in pay. With the onset of the war, federal controls multiplied. Selective price controls were imposed and crucial industries taken over and operated by the government. The majority of the items affected by price controls were industrial commodities and raw materials that the government required in great quantities. For consumers the most important price control was the extensive imposition of rent controls, especially in the industrial and port cities where war-induced activity grew sharply. Excess-profits taxes helped reduce property incomes while the war and government encouragement of unions and collective bargaining pushed up real wage rates, particularly for unskilled nonfarm workers. The authorities controlling wages allowed even larger increases “for the more poorly paid workers, based on the principle that everyone should be able to afford a family budget representing a ‘minimum of health and decency.’” On the other hand, salaried workers, whose incomes were normally higher, saw their real incomes fall because “they did not engage in collective bargaining, and there were no boards to alter their compensation as the cost of living rose.” The result was a rise in equality as lower paid workers gained income shares at the expense of salaried workers and individuals whose incomes came primarily from investments.

12 Williamson and Lindert, American Inequality, chap. 12.
13 Soltow, “Evidence.”
14 The inverse-Pareto slope coefficient declined from an average of 0.70 in the period from 1866 through 1871 to 0.685 in the period from 1913 through 1916. The coefficient fell to 0.61 in 1919 and to 0.60 in 1920. On average it was 0.55 in the 1920 to 1924 period and 0.67 in the 1925 to 1929 period. Note that the 1925 through 1929 average is almost back to the 1913 through 1916 average. (Soltow, “Evidence,” p. 282.)
17 Higgs, Crisis, pp. 116–21.
18 See Higgs, Crisis, chap. 7, especially pp. 141 and 158; and, see Rockoff, Drastic Measures, chap. 3. For a more extensive discussion of rent controls during the First World War see Schaub, “Regulation.” I am indebted to an anonymous referee for suggesting these additional references.
20 Ibid., p. 74.
21 Ibid., p. 76.
22 Hughes and Cain, American Economic History, p. 416.
As Williamson has noted, "Presumably, these forces are reversed as the economy returns to 'normalcy.'" The changes in labor demands led to a fall in unskilled wages and a rise in skilled wages, a trend augmented by the postwar decline in labor-union influence. It was 1927 before real unskilled wages reached the level they achieved in 1920, while the real wages of skilled workers rose 11 percent between 1920 and 1927. The removal or reduction in other controls allowed property incomes to rise. The question is how long did it take to return to "normalcy," and how much of the rise in inequality in the twenties can be explained by this? Soltow's data suggest that this may have been completed by the mid-twenties because his measure of upper-income inequality for the 1925 through 1929 period was essentially the same as for the 1913 through 1916 period.

Capital Gains and the Stock Market Boom

The other major influence in the twenties was the great stock market boom of the last half of the decade. Clearly this conferred large increases in disposable income on those who realized capital gains during the boom. It was possible to make some crude calculations to measure the contribution of the excess of gains over losses on asset sales toward inequality changes in the twenties. To do this the shares of disposable income were re-estimated completely excluding the adjustment for the excess of gains over losses on the sales of assets.

The estimates excluding realized capital gains show significantly different trends. The gains in income shares for the top percentiles are all concentrated in the 1919 to 1922 period. Between 1923 and 1929 the share of income for the top 1 percent tends to fall slightly though the trend is not significant. The loss in income share for the lower 93 percent is concentrated in the 1919 to 1922 period, and there is no significant trend in their share between 1923 and 1929.

Gini coefficients are the best overall measure of how the exclusion of realized capital gains alters changes in the size distribution of income. These coefficients for income shares of the nonfarm population when realized capital gains are excluded are presented in Table 1. Through 1922 the Gini coefficients are quite similar, though the estimates excluding realized capital gains show somewhat less inequality. In 1923 and 1924 the difference grows considerably. Again in 1925 through 1927 there is a rise to a larger difference. Finally, in 1928 and 1929 the difference is the largest as the estimates excluding realized capital gains are 20.28 and 18.31 percentage points smaller. The Gini coefficients for the estimates that exclude realized capital gains show a negative trend from 1923 through 1929, however it is not a significant trend.

Explaining Changes in the Size Distribution of Income

These examinations suggest the following preliminary explanation for the increased inequality in the size distribution of income in the twenties. The general conclusion is that there would have been some increase in inequality in the return to "normalcy" following the First World War. The forced move to more equality brought about by wartime controls was not sustainable in a freer environment. Soltow's analysis using unadjusted incomes of the upper-income percentiles suggested this was accomplished.

---

26 See Smiley, "New Estimates," table 3, page 227, for the data from which these changes were calculated.
by around 1925. This analysis of Gini coefficients using the new income-share estimates including and excluding realized capital gains suggests that this return to normalcy was largely completed by 1922 because the Gini coefficients are quite similar in movement and level from 1919 through 1922.

From 1923 on the Gini coefficients diverge and there appear to be three stages in this divergence brought on by the realization of capital gains in the securities market boom of the twenties. From 1919 to 1922 the Gini coefficient excluding realized capital gains averaged about 3.65 percentage points less than the Gini including realized capital gains. In the initial stages of the securities-market boom in 1923 and 1924 the difference jumped to 8.17 percentage points as capital gains began to be realized by higher income individuals. As the securities-market boom gathered force in 1925, 1926, and 1927 the difference rose to 15.04 percentage points, and in the final stages of the boom in 1928 and 1929 the difference rose to 19.3 percentage points. Excluding realized capital gains eliminates the increases in income inequality that occurred after 1922.

CONCLUSIONS

Simon Kuznets's study of the income shares of upper-income groups in the interwar period is and will remain a masterful study. However, it now seems clear that his estimates of income shares during the twenties are biased because he failed to adjust for the changes in tax-avoidance behavior of taxpayers during this decade. The rise in income inequality was less pronounced than Kuznets found. In particular, the lower 93 percent of the nonfarm population did not experience decreases in their real and nominal per capita incomes after the recovery from the 1920/21 depression. Though their real income growth was much slower, it did grow along with the real per capita incomes of the nonfarm upper-income percentiles.

The increases in inequality that appear in the revised estimates for the 1920s seem to have been driven by two factors. First, there was an initial rise in inequality during the return to "normalcy" after the First World War. This seems to have been completed by the end of 1922. From 1923 on, the rising inequality was primarily driven by the booming stock market as the excess of gains over losses on asset sales appears to explain most of the differences in real per capita income growth.

It is necessary to point out that these new estimates are still far from satisfactory. They fail to consider the level of and changes in income inequality among the lower 93 percent of the nonfarm population and they fail to adjust for income status to get a truer picture of how the distribution of income changed over time. Both of these points reflect the data limitations of this period.

Williamson and Lindert have considered income inequality among the large nonfarm lower income group. They argue that changes in the ratio of skilled to unskilled wages are good proxies for changes in inequality among the lower 93 percent of the nonfarm population—the group for which no income data are available. They found that this ratio positively tracked Kuznets's top 1 percent share and it would likely have a similar relationship to the new estimates of the income share of the top 1 percent. This suggests that inequality among the nonfarm lower 93 percent increased during the 1920s, though we

27 Williamson and Lindert, American Inequality, pp. 80–82.
28 Williamson and Lindert found a significantly positive coefficient on the wage ratio variable in the regression of the Kuznets's top 1 percent share on the wage ratio and on unemployment, though much of the variation in the top 1 percent's share was unexplained by the regression (pp. 80–81). Though the new estimates of the income share of the top 1 percent rise less than Kuznets's estimates, the patterns are similar enough that the positive relationship between the top 1 percent's share and the wage ratio would almost certainly exist if the new estimates are used.
cannot suggest just how much it increased. It is likely, therefore, that the Gini coefficients for both Kuznets's original income-share estimates and the new income-share estimates to some extent understate the rise in income inequality during the 1920s.

These estimates, as did Kuznets's original estimates, fail to adjust for income status. At any point in time each percentile will contain migrants, income recipients with temporarily higher or lower incomes compared to their normal incomes. To assess changes in income shares (or changes in the income distribution) over time, income recipients should be classified by income status for relatively long periods, say at least five to ten years. In this way each percentile would contain only residents and no migrants. Then changes in income shares over time would actually show changes in income shares for income recipients because the same people would be in each income status group at each point in time.

Kuznets had no data to construct income status groups, however he could draw upon Frank Hanna's examination of Wisconsin incomes. For Wisconsin Hanna found that when families were classified by income for three years rather than single years the share of the top 5 percent was about 3.3 percentage points lower. Kuznets conjectured that, "in passing from a distribution by size of income in a given year to a distribution by income status for five years to a decade, the share of the top 1 percent of the former should be cut by about a fifth; and that of the top 5 percent, about a seventh." Therefore both Kuznets's original estimates and the new estimates overstate the degree of income inequality during the 1920s.

However, Kuznets made no conjectures on how, or if, an analysis of income status groups would effect changes in the distribution of income during the 1920s. Such evidence is not available. Evidence from the last 30 years suggests that movement between percentile groups over a ten-year period, or even a single year period, is very large. Unfortunately we do not know whether income mobility in the 1920s was similar to that in the 1970s and 1980s. Thus, though we can say that these estimates overstate the level of income inequality in the 1920s, we cannot say whether the correction would show smaller or larger increases in inequality during the decade.

Finally, it must be admitted that these estimates are crude, and one hopes that they will be improved in the future. However, even with the crudity of these new income-share estimates, it now appears that the gains of the prosperous 1920s were more broadly based than Kuznets's share estimates had suggested. The mass of the nonfarm population did see increases in their real incomes and the rise in inequality was less pronounced than previously thought.

---

29 Hanna, Analysis.
30 Ibid., tables 11 and 12.
31 Kuznets, Shares, p. 140.
32 For a summary of this see Smiley, American Economy, pp. 404–12.
33 The estimates would be improved if there were better information on how the holdings of tax-exempt state and municipal securities by the upper-income groups changed as tax rates changed in the twenties. If tax rates had been constant at lower rates, this surely would have had real effects on the economy. Interest rates, savings, and the allocation of investment spending might well have been different. This would likely have had effects on national income and a more sophisticated analysis would take this into effect.
REFERENCES


