

NBER WORKING PAPER SERIES  
ON  
HISTORICAL FACTORS IN LONG RUN GROWTH

WEALTH INEQUALITY TRENDS IN INDUSTRIALIZING  
NEW ENGLAND: NEW EVIDENCE AND  
TESTS OF COMPETING HYPOTHESES

Richard H. Steckel  
Carolyn M. Moehling

Historical Paper **122**

NATIONAL BUREAU OF ECONOMIC RESEARCH  
1050 Massachusetts Avenue  
Cambridge, MA 02138  
February 2000

Any opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

© 2000 by Richard H. Steckel and Carolyn M. Moehling. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Wealth Inequality Trends in Industrializing New England:  
New Evidence and Tests of Competing Hypotheses  
Richard H. Steckel and Carolyn M. Moehling  
NBER Historical Paper No. 122  
February 2000  
Development of the American Economy

**ABSTRACT**

This paper assembles new data and new methods for studying wealth inequality trends in industrializing America. Records of household heads from the census matched with real and personal property tax records for Massachusetts reveal that the Theil entropy measure of inequality approximately doubled over the period from 1820 to 1910, a gain that was divided about evenly between the antebellum and the postbellum periods. A surge between 1870 and 1900 dominated the growth in inequality following the Civil War. Decompositions of changes in the Theil entropy measure reveal that during both periods, inequality was increasing due to the shift of the population out of rural areas and agriculture into urban areas where wealth was less equally distributed. But the increases in inequality were also due to increasing inequality within population groups. Between 1870 and 1910, inequality was growing within occupations, age groups, and the native-born population. Proposed labor market explanations, including sectoral shift that led to higher wages in non-agricultural relative to agricultural sectors, biased technological change, and immigration are inconsistent with the fact that inequality between occupational groups was declining in the last decades of the century. Wealth accumulation patterns by age are also inconsistent with the hypothesis of child default on responsibilities for old age care, at least during the second half of the nineteenth century. To explain the salient facts, we are led to propose a new explanation based on luck, rents and entrepreneurship.

Richard H. Steckel  
Economics Department  
Ohio State University  
Columbus, OH 43210  
and NBER  
steckel.1@osu.edu

Carolyn M. Moehling  
Department of Economics  
Yale University  
New Haven, CT 06520  
and NBER  
carolyn.moehling@yale.edu

Economic, social, and political historians have long studied the distributions of income and wealth because inequality affects economic growth via savings, investment, and incentives for taking entrepreneurial risks. The degree of inequality within and across groups also identifies the distribution of economic rewards, which is an important factor governing influence in social and political processes.

Patterns of inequality from wealth or income data are reasonably well established for three eras in American history: the late 1700s, the mid-nineteenth century, and from approximately 1920 onward.<sup>1</sup> After surveying the available evidence two decades ago, Jeffrey Williamson and Peter Lindert concluded that inequality increased from the late 1700s through the second quarter of the twentieth century and declined thereafter.<sup>2</sup> In their judgment, much of the nineteenth-century trend toward greater inequality occurred with the onset of industrialization, in the four decades prior to the Civil War. They described the period between the Civil War and the Great Depression as an “uneven plateau” during which inequality remained high but exhibited no long-term trend. They recognized, however, that additional evidence on the distribution of wealth and income during this period was sorely needed.<sup>3</sup>

---

<sup>1</sup> Simon Kuznets, Shares, and Robert Lampman, Share, pioneered the study of inequality trends in this century using federal income tax records and federal estate tax data. Lee Soltow, Men and Wealth, and others anchored inequality estimates for the mid-nineteenth century in household wealth reported by the federal censuses of 1850, 1860, and 1870. Soltow, Distribution, also examined inequality using real estate appraisals from the First Direct Tax of 1798. Alice Hanson Jones, Wealth, and the Mains, “Inequality,” and Social, developed and refined probate records as a source for the late Colonial and early national period. For a recent survey of the literature see Lindert, “Three Centuries.”

<sup>2</sup> Williamson and Lindert, American Inequality.

<sup>3</sup> Ibid, p. 43. Lee Soltow, “Inequalities,” however, has argued that inequality did not increase significantly during the nineteenth century.

This paper develops and analyzes new data on long-term trends in inequality from 1820 to 1910—nearly 10,000 male household heads listed in the manuscript schedules of the census matched with real and personal property tax records from Massachusetts. This state was by far the largest within industrializing New England. It was also at or near the center of industrial change in the region, which included the structural shift out of agriculture, the growth of the textile industry, rapid urbanization, immigration, and out-migration—all factors thought to have had effects on the distribution of wealth in the industrializing Northeast.

We investigate trends in wealth inequality using measures widely used by economists in the inequality literature: the shares of wealth held by the top fractiles of the wealth distribution, the Gini coefficient, and the Theil entropy measure. The Theil entropy measure is particularly useful because its changes over time can be decomposed into components representing the effects of shifting population shares, changes in the relative mean wealth of different population sub-groups, and changes in within-group inequality.

## **Data**

Although a separate paper has been published on our data source, for the convenience of readers we begin with a brief description of the evidence, which is new to the literature on long-term trends in inequality.<sup>4</sup> Real and personal property taxes formed the backbone of state and local tax revenues until income and sales taxes were introduced in the twentieth century.

According to Richard T. Ely, the antebellum period “witnessed the complete establishment of the American system of state and local taxation. The distinguishing feature is ... the taxation of all

---

<sup>4</sup> See Steckel, “Census Manuscript Schedules,” which gives details on sampling procedures, additional characteristics of the sample, detailed definitions of occupations, collection of taxes, and comparisons with wealth reported by the censuses of 1850, 1860, and 1870.

property, moveable and immovable, visible and invisible, real and personal, as we say in America, at one uniform rate.”<sup>5</sup> The similarity of the tax bases in the states is made clear from digests of taxation.<sup>6</sup> Legislation created assessors, boards of equalization, and other machinery to collect taxes. Reevaluations were made at irregular intervals but ownership lists of taxable property typically were prepared every year, not merely at the time of reevaluations. State governments operated through local authorities (county or town) to conduct the assessments and maintain the tax lists of individuals. In 1796 the list of ratable property in Massachusetts was “so long as to include almost everything.”<sup>7</sup> All real and personal property not specially exempted was subject to taxation. Real estate included land and buildings, and personal estate included goods, chattels, money and effects (wherever they were); ships; money at interest; public stocks and securities; stocks in turnpikes, bridges, and moneyed corporations, in or out of state. Property exempted from taxation included household furniture not exceeding \$1000 in value, wearing apparel, farming utensils, and mechanics' tools up to the value of \$300.<sup>8</sup>

The manuscript schedules of federal population census became more elaborate and comprehensive during the nineteenth century.<sup>9</sup> Through the census of 1840, census schedules listed the name of each household head and reported the number of household members in different demographic groups defined by sex, age, and race. Starting with the 1850 census, the schedules listed each individual by name and reported each individual's age, sex, color, and

---

<sup>5</sup> Ely, Taxation, p. 131.

<sup>6</sup> Street, Digest and Plehn, Revenue Systems.

<sup>7</sup> Ely, Taxation, p. 138.

<sup>8</sup> Bullock, “Taxation;” Nichols, Taxation.

<sup>9</sup> Wright, History.

other characteristics such as occupation. The 1850 census listed occupations for males over 15, the value of real estate owned, and place of birth. The 1860 schedules added personal property and occupations for females, and the 1870 schedules added the nativity of one's father and mother. Beginning in 1880, the schedules dropped wealth but added questions on unemployment, relationship to the household head, and the birthplaces of one's parents.<sup>10</sup> The schedules for 1900 and 1910 added questions on the year of immigration to the U.S., ability to speak English, and home ownership.

Following a survey of the available tax records, samples were taken from localities that had a complete set of records over the period 1820 to 1910. The sample for Massachusetts includes Boston, Salem, Lexington, Westminster, and Sturbridge. In each census year, approximately 1,200 households were randomly chosen from each of the urban and the rural areas of these townships. All information available for each household was recorded from the census manuscript schedules.

The tax records were maintained in alphabetical order of the taxpayer by ward in cities or by town or township in rural areas. The census manuscript schedules were alphabetized accordingly to facilitate the search for a match. If a household head was not found in the tax records, it was assumed that he or she had no taxable property. This assumption may lead to errors in tabulating wealth in cases of garbled names or where individuals moved between the dates of the census and the tax enumeration. Since matches were sometimes ambiguous, a coding procedure was devised to rate the confidence of the match, with categories of exact match, nearly exact match, probable match, improbable match, and duplicate (two or more

---

<sup>10</sup> Unfortunately, nearly all the 1890 schedules were destroyed by fire.

people with the same name). The last two categories, which amounted to 2.1 to 4.1 per cent of the sample (depending upon census year), were omitted from calculations.

Property tax records matched with census manuscript schedules provide a valuable, new source to measure and analyze long-term trends in inequality. There were strong incentives for accuracy in the tax assessment system: tax collectors sought to obtain all revenue permitted by law and property owners could appeal unfair assessments. In our judgment, taxable wealth data are more reliable than the self-reported wealth data collected in the 1850, 1860, and 1870 censuses. The 1860 instructions to the enumerators concerning the reporting of personal property wealth state: “Exact accuracy may not be arrived at, but all persons should be encouraged to give a *near and prompt* estimate for your information” (emphasis added). The reported wealth levels are clustered on multiples of 100, indicating a strong tendency for rounding in the self-reports. Timothy G. Conley and David W. Galenson have also pointed out that the census wealth data appear to be censored, but the point of censoring is uncertain and may have varied across enumerators.<sup>11</sup>

Richard Steckel uses scatter diagrams and regressions to compare census wealth with taxable wealth for the 1850, 1860, and 1870.<sup>12</sup> In the case of discrepancies, census wealth often exceeded taxable wealth, but the differences were not systematically associated with socioeconomic variables, such as occupation or age, that were reported by the census.<sup>13</sup> There are several plausible explanations for the differences, including assessments below market value,

---

<sup>11</sup> Conley and Galenson, “Quantile Regression Analysis.”

<sup>12</sup> Steckel, “Census Manuscript Schedules.”

<sup>13</sup> Taxable wealth was systematically lower for widows, who received favorable tax treatment.

exemptions, and inclusion of property owned by the spouse or children in census wealth. Gini coefficients were larger in the tax records, but the differences were less than 0.06.

In this paper, we limit our analysis to the distribution of wealth across male household heads.<sup>14</sup> Table 1 provides economic and demographic characteristics of the samples. The data in each sample have been weighted so that the share urban in the sample equals the share urban in the state for that census year. During the sample period, the population became highly urbanized, with the share living in cities and towns increasing from 22.8 percent in 1820 to 89.0 percent in 1910. Over the century, the population aged, and the share of heads age 50 and older grew. Reflecting the immigration of the period, the share of foreign-born increased between 1850 and 1910. The decline of farming in Massachusetts is evident in the compositional shift of the occupational structure away from agriculture.<sup>15</sup>

The percent of the household heads without a match more than doubled between 1820 and 1910 (from 34.2 percent to 73.6 percent), raising the question of whether match failures were caused by something other than growing inequality. Perhaps fewer individuals were matched because migration rates increased or growth in city size complicated the maintenance of accurate records. One test of this possibility is provided by data on the number of individuals assessed for

---

<sup>14</sup> The taxable wealth of females, who comprised about 10 percent of the initial sample, will be studied in a separate paper.

<sup>15</sup> The characteristics and trends in characteristics of our sample are similar to those of the samples of household heads from the IPUMS data for Massachusetts, Connecticut and Rhode Island. The trends observed for the samples of household heads mirror those observed in the IPUMS data for the adult male population and therefore reflect general population changes rather than changes in the patterns of headship. Headship rates for all age groups except ages 70 and older did decline slightly between 1850 and 1910, which means that our sample of household heads is likely slightly “more selective” in 1910 compared with 1850. Since most adult males who were not household heads likely had little or no wealth, this implies that we may be understating the increase in inequality over the period.



only the poll tax (a head tax) relative to the total number of polls assessed. If slippage or disagreement between the census and the tax records was small, then this ratio should approximately equal the share of household heads not found in the tax records. This is indeed the case. The ratio of the number assessed for the poll tax only to total male polls assessed was 67.2 percent in 1886, 70.8 per cent in 1890, 72.0 percent in 1900 and 77.9 percent in 1910.<sup>16</sup> These levels and the trend approximately agree with those in Table 1 for the percent with no match near the turn of the century.

Table 2 shows that the majority of wealth was held in real estate. Over time, the very rich (top 1%) held a somewhat increasing share of their total wealth in this form, which raises the possibility that some financial wealth may have been omitted from the tax base. According to Robert Lampman, estate tax data show that real estate declined as a share of total assets as wealth increased, and that in 1922 about 28 percent of economic estate (total gross estate minus debts, which is roughly comparable to taxable property in Massachusetts) was held in real estate.<sup>17</sup> On the other hand, the estate tax exemption in 1922 was \$50,000, which confined the estate tax data to the very rich (about 1 percent of all deaths), a group likely to have held less wealth in real estate. Although it would be desirable to compare our taxable wealth data with other measures of total assets in the late nineteenth century, it is apparent that any downward bias in measured inequality (arising from omitted financial wealth) cannot have been very large. The Gini coefficients (discussed in connection with Table 3) near the turn of the century were already extraordinarily high (about 0.90), and including any financial wealth that may have been omitted would only have made them greater. In conclusion, we also note that if growth in

---

<sup>16</sup> The results are taken from Commonwealth of Massachusetts, Aggregate of Polls.

<sup>17</sup> Lampman, Share, p. 158.

financial wealth and its omission from the tax base was a problem, then our measures understate the extent, but not the direction of inequality change in the late nineteenth and early twentieth centuries.

### **Trends in Wealth Inequality, 1820 to 1910**

We use three measures of inequality to examine long-run changes in the distribution of wealth: the shares of total wealth held by the top fractiles of the wealth distribution, the Gini coefficient, and the Theil entropy measure. Previous research on historical wealth distributions has focused on the first two of these measures. The shares of wealth held by the top 20 percent or the top 5 percent of the wealth distribution are straightforward and easy to calculate, but they do not capture the degree of dispersion within the top fractiles and ignore the lower fractiles of the wealth distribution. The Gini coefficient is an index measure based on the average absolute difference in wealth levels between all pairs of individuals or households. The Gini also has the more intuitive interpretation as twice the area between the Lorenz curve and the diagonal representing the case of “perfect equality” when all individuals have the same level of wealth.

Like the Gini, the Theil entropy measure is an index measure based on the entire wealth distribution. The Theil measure is given by the following equation:

$$(1) \quad T = \frac{1}{n} \sum_{i=1}^n \frac{w_i}{\mu} \ln \left( \frac{w_i}{\mu} \right)$$

where  $n$  represents the number of observations,  $w_i$  represents the wealth of individual  $i$ ,  $\mu$  represents the full sample mean wealth, and  $[0 \ln(0)]$  is taken to be zero.<sup>18</sup> In the case of “perfect equality” when all individuals have the same level of wealth, the Theil measure, like the Gini,

---

<sup>18</sup> Foster, “Inequality,” p. 55.

equals zero. In the case of “perfect inequality” when one individual owns all of the society’s wealth, the Theil measure equals  $[\ln(n)]$ . Although this means that the maximum value of the Theil measure varies with sample size, the rapidly diminishing slope of the natural log function makes this of little practical importance for samples, such as ours, that are fairly large and approximately the same size.

For all of the calculated inequality measures, we use bootstrap methods to estimate approximate standard errors, construct confidence intervals, and perform hypothesis tests.<sup>19</sup> For each sample of size  $n$ , we construct 1000 re-samples of size  $n$  by random draws with replacement from the original sample. Following Mills and Zandvakili, we use the “percentile method” to construct confidence intervals, calculating tail probabilities directly from the bootstrapped distribution.<sup>20</sup>

Table 3 reports the aggregate measures of wealth inequality for the samples of male household heads. The three types of measures are highly correlated: increases in the shares of wealth of the top 20, 5, and 1 percent of the wealth distribution correspond to increases in the Gini coefficient and the Theil entropy measure. The correlation between the Gini and the Theil measure is 0.97. The choice of inequality measure, therefore, has little impact on the observed trends in wealth inequality.

Figure 1 charts the Gini coefficients and Theil entropy measures. The dashed lines represent the 95 percent confidence intervals constructed from bootstrap analysis. Wealth

---

<sup>19</sup> Asymptotic approximations of the variances of the Gini coefficient and the Theil entropy measure do exist, but little is known of their small sample properties. Statistical inference based on bootstrap methods has been shown to be superior to asymptotic approximations both on theoretical grounds and in a variety of applications. See Mills and Zandvakili, “Statistical Inference.”

<sup>20</sup> Mills and Zandvakili, “Statistical Inference.”

inequality was substantially higher in Massachusetts in 1910 than it had been in 1820. The rise in inequality, though, was not steady. Both the Gini and the Theil measure indicate that wealth inequality grew sharply between 1820 and 1850, leveled off between 1850 and 1870, and then began a steady increase to 1900. The increases in both measures between 1820 and 1850, and 1870 and 1910 are statistically significant at the 5% level.<sup>21</sup> The data, therefore, reveal two periods of increasing inequality: the four decades prior to the Civil War previously noted by Williamson and Lindert, and 1870 to 1910. As seen in Table 3, during both of these periods, wealth became increasingly concentrated in the top fractiles of the wealth distribution. The share of wealth held by the top 20 percent rose from 72 to 86 percent between 1820 and 1850 and from 90 to 98 percent between 1870 and 1910.

## Decomposing the trends

### *Decompositions of the Theil Entropy Measure*

To gain insights into the forces underlying the increases in inequality observed in Table 3 and Figure 1, we consider the decompositions of the Theil entropy measure and its changes over time. The additivity of the Theil measure allows it to be decomposed into components representing inequality arising from differences in wealth levels *between* population sub-groups and inequality arising from variation in wealth levels *within* those groups. For any exhaustive collection of mutually exclusive subsets of observations labeled  $\{1, 2, \dots, G\}$ , this measure can be re-written as follows:

$$(2) \quad T = \sum_{g=1}^G \frac{n_g \mu_g}{n\mu} T_g + \sum_{g=1}^G \frac{n_g \mu_g}{n\mu} \ln \left( \frac{\mu_g}{\mu} \right)$$

---

<sup>21</sup> These tests were conducted by using bootstrap analysis to calculate approximate standard errors and confidence intervals for the *difference* between periods.

where  $n_g$  represents the number of observations in sub-group  $g$ ,  $\mu_g$  represents the mean wealth of sub-group  $g$ , and  $T_g$  represents the measure in equation (1) calculated for sub-group  $g$ .<sup>22</sup> The first term on the right hand side of equation (2) is the weighted sum of the Theil entropy measures for the sub-group wealth distributions where the weights are the sub-group shares of total wealth. This term represents the component of measured inequality due to inequality in the distribution of wealth *within* population sub-groups. The second term is simply the Theil entropy measure of equation (1) calculated from a wealth distribution in which each person is assigned the mean wealth of their sub-group, and, therefore, represents the component of measured inequality due to inequality in the distribution of wealth *between* population sub-groups.<sup>23</sup>

The decomposition of aggregate measured inequality into its within-group and between-group components can provide important clues as to the sources of inequality in a society. For instance, if a substantial fraction of aggregate inequality is due to inequality *between* occupation categories, this suggests that an important source of inequality is systematic variation in the returns to particular skills.

More importantly, for the study of long-term trends in inequality, changes in the Theil entropy measure may be decomposed into three components: (i) changes in the population

---

<sup>22</sup> Foster, “Inequality,” p. 56. Showing that the right-hand-side of equation (2) is equal to the right-hand-side of equation (1) is straightforward. Note that the sub-group level Theil measure,  $T_g$ , can be expressed as:

$$T_g = \frac{1}{n_g} \sum_{i=1}^{n_g} \frac{w_i}{\mu_g} \ln \left( \frac{w_i}{\mu_g} \right)$$

To derive equation (1), simply substitute this expression into (2) and rearrange terms so that the summations are over  $i=1 \dots n$  instead of  $g=1, \dots, G$ .

<sup>23</sup> The Theil entropy measure is a member of the Generalized entropy class of measures. Cowell and Kuga, “Additivity,” and Cowell, “The Structure,” have shown that only measures in this class satisfy a set of axioms considered desirable in inequality measures: a weak principle of transfers, general decomposability, and complete scale invariance.

shares of sub-groups ( $n_g/n$ ); (ii) changes in the relative mean wealth of subgroups ( $\mu_g/\mu$ ), and (iii) changes in the dispersion of wealth within subgroups ( $T_g$ ). The contributions of each of these elements can be calculated as follows:

$$(3) \quad \Delta T_n^{t,s} = \Delta T_{\text{within},n}^{t,s} + \Delta T_{\text{between},n}^{t,s} = \sum_{g=1}^G \left[ \left( \frac{n_g^t}{n^t} - \frac{n_g^s}{n^s} \right) \frac{\mu_g^t}{\mu^t} T_g^t \right] + \sum_{g=1}^G \left( \frac{n_g^t}{n^t} - \frac{n_g^s}{n^s} \right) \frac{\mu_g^t}{\mu^t} \ln \left( \frac{\mu_g^t}{\mu^t} \right)$$

$$\Delta T_\mu^{t,s} = \Delta T_{\text{within},\mu}^{t,s} + \Delta T_{\text{between},\mu}^{t,s} = \sum_{g=1}^G \left[ \left( \frac{\mu_g^t}{\mu^t} - \frac{\mu_g^s}{\mu^s} \right) \frac{n_g^s}{n^s} T_g^t + \sum_{g=1}^G \left[ \frac{\mu_g^t}{\mu^t} \ln \left( \frac{\mu_g^t}{\mu^t} \right) - \frac{\mu_g^s}{\mu^s} \ln \left( \frac{\mu_g^s}{\mu^s} \right) \right] \frac{n_g^s}{n^s} \right]$$

$$\Delta T_T^{t,s} = \sum_{g=1}^G \frac{n_g^s \mu_g^s}{n^s \mu^s} (T_g^t - T_g^s)$$

This decomposition provides important insights into the sources of change and allows us to evaluate some of the proposed explanations of rising inequality trends in the nineteenth century to which we now turn.

### ***Proposed Explanations and Their Testable Implications***

Suggested explanations of rising inequality in the nineteenth century can be grouped into four categories: composition effects, structural shift, labor market changes, and the rise of life-cycle behavior. We give a brief synopsis of important hypotheses within each category and discuss their testable implications.

The most straightforward hypotheses attribute changes in inequality to changes in the composition of the population. In his well-know article linking rising inequality with industrialization, Simon Kuznets suggested that urbanization contributed to the process because inequality was greater in urban compared with rural areas.<sup>24</sup> Immigration may have acted on

---

<sup>24</sup> Kuznets, “Economic Growth,” pp. 7-8.

inequality in much the same way as urbanization. Immigrants were poor relative to the native born and the share of immigrants in the total population rose in two waves: the 1830s to the 1850s and from the 1860s to the early twentieth century.<sup>25</sup> These waves coincide with spurts of inequality growth found in our data.

Another explanation of changes in inequality in the nineteenth century focuses on the changes in the population age structure. Age is an important determinant of wealth. Studies for the nineteenth century show that wealth tended to peak for people in the mid or late 50s. Given this phenomenon, Jeremy Atack and Fred Bateman have noted that wealth would be more unequally distributed in a society skewed toward younger ages.<sup>26</sup> Given the decline in fertility in the nineteenth century and its impact on raising the average age of the population, however, this hypothesis cannot explain the *rise* in inequality we observe in the Massachusetts data. Yet, it is relevant to contemplate ways that the aging of the population could have dampened the effects of other forces leading to inequality.

More interesting theories of increasing inequality in the nineteenth century propose widening gaps *between* population subgroups. Kuznets suggested that the gap between agricultural and industrial workers may have grown. Since the income elasticity of demand for food is less than one, economic growth is presumed to induce a rise in the demand for industrial products and services and a decline in the demand for food. This structural shift leads to stagnating wages in agriculture relative to the other sectors, which contributes to rising inequality

---

<sup>25</sup> See Williamson and Lindert, American Inequality, p. 209.

<sup>26</sup> Atack and Bateman, "Egalitarianism."

over time.<sup>27</sup> This explanation also suggests a widening gap between rural and urban wealth. If correct, much of the rise in the Theil measure of inequality should be attributable to greater inequality between the agricultural and non-agricultural subgroups and between the urban and rural population.

Several investigators have argued that the skill-premium rose during the nineteenth century. Peter Lindert and Jeffrey Williamson suggest this followed from capital accumulation and biased technological change, which increased the demand for skilled relative to unskilled labor and led to “wage stretching” or a widening in gap in pay.<sup>28</sup> Waves of immigration also increased the supply of unskilled relative to skilled labor, and may have depressed the wages of the unskilled relative to the skilled. In his “frontier thesis,” Frederick Jackson Turner argued that the west acted as a safety valve that absorbed large quantities of unskilled labor originating in the east, which buoyed wages of the unskilled in eastern labor markets. When the frontier closed near the end of the nineteenth century, unskilled eastern workers lost any earnings advantage conveyed by this safety valve. All three of these mechanisms have the same testable implications for the decomposition of changes in the Theil inequality measure. Namely, they all predict a rise in between group inequality when the wealth holders are organized by occupations.

Roger Ransom and Richard Sutch have proposed a very different mechanism of change. They argue that growing inequality in the nineteenth century was the result of new life-cycle patterns of wealth accumulation that were connected to declines in fertility.<sup>29</sup> In their view, the initiators of change were children who moved West and defaulted on obligations to provide old-

---

<sup>27</sup> Kuznets, “Economic Growth,” pp. 7-8.

<sup>28</sup> Lindert and Williamson, American Inequality.

<sup>29</sup> Ransom and Sutch, “Two Strategies;” Sutch, “All Things Reconsidered.”



age care for their parents. This behavior led subsequent generations of parents to have fewer children and instead accumulate more wealth, which was then spent on old-age care in lieu of direct support by their children. These changes in behavior led to a growing concentration of wealth by age. Specifically, other things being equal, the mean wealth of household heads in their prime saving years (40s and 50s) should have risen relative to the mean wealth of other age groups.

### ***Decomposition Results***

Table 4 presents the decompositions of the changes in the Theil entropy measure from 1820 to 1850 and 1870 to 1910 – the two periods of increasing inequality. These data indicate that the increasing inequality was due to both shifts in population shares and increases in within-group inequality.

Just as Kuznets proposed, the shift of the population into urban areas led to greater inequality. Between 1820 and 1850, the urban population in Massachusetts rose from less than a quarter to half of the state's population. This change alone would have led to a greater than 50 percent increase in measured inequality, holding inequality within and between urban and rural populations unchanged. Likewise, the rise in the urban population from 67 to 89 percent between 1870 and 1910 accounts for a 22 percent rise in aggregate inequality.

The shift of the population out of agricultural occupations between 1870 and 1910, which was strongly related to the process of urbanization, also can account for a substantial increase in inequality. In fact, had inequality between and within occupation groups been unchanged between 1870 and 1910, measured inequality would have increased by 41 percent just because of changes in the occupational distribution of household heads in Massachusetts.

The other composition effects were small. As expected, the aging of the population did have a slight negative effect on aggregate inequality. Surprisingly, however, the increase in the foreign-born population also had a negative effect on aggregate inequality.

While the composition hypotheses of changing inequality find support in Table 4, the hypotheses predicting widening gaps between groups do not. In fact, the data indicate that wealth gaps between occupation groups and between the urban and rural populations were decreasing rather than increasing. The prominence of the hypotheses predicting growing disparities between groups motivated us to test these results further by estimating tobit regressions for the log of wealth for the years 1850, 1870, and 1910. The explanatory variables included in these regressions are indicators of age, urban residence, race, nativity, illiteracy, and occupation.<sup>30</sup> Table 5 presents the estimated coefficients and the predicted effects on observed wealth. The predicted effects were calculated as the ratio of the expected *observed* wealth of individuals with a particular characteristic to the expected *observed* wealth of the baseline individual.<sup>31</sup> So for instance, the predicted effect reported for the 30 to 39 year-old age group in 1870 is 3.730. This indicates that controlling for other factors, the wealth of males age 30 to 39 was almost 4 times that of males 20 to 29.

The data in Table 5 indicate that inequality between age groups did not increase between 1870 and 1910. The data in Table 5, in fact, indicate that differences in wealth levels between age groups fell between 1870 and 1910. But differences in wealth levels between age groups did

---

<sup>30</sup> The dependent variable was defined as:  $\ln(\text{wealth} + \$1)$ .

<sup>31</sup> The expected observed wealth of different groups are calculated using the following formula:

$$E[\ln(w_i + 1)] = \Phi_i(x_i'\beta/\sigma) x_i'\beta + \sigma\phi_i(x_i'\beta/\sigma)$$

grow between 1850 and 1870, even though aggregate inequality during this period was fairly stable. Between age-group differentials were also larger in 1910 than in 1850. However, the changes in these differentials are not consistent with the predictions of Ransom and Sutch's hypothesis. Although all groups gained relative to the baseline group of 20 to 29 year olds, the biggest gains were made by males in their 60s and above, the group that Ransom and Sutch's hypothesis predicts should have been spending their accumulated assets. The gains made by men in their 40s and 50s were fairly small.

The results regarding occupation differences in wealth are also intriguing. Table 5 indicates that these differences were not only not growing between 1850 and 1910, but, for the most part, were falling over this period. The relative wealth of white collar workers fell substantially between 1850 and 1910. By 1910, farmers had overtaken white collar workers as the group with the highest mean wealth. More interesting given the hypotheses predicting wage-stretching during this period, the gap between the skilled and the unskilled also fell substantially. In 1850, the predicted ratio of skilled to unskilled workers wealth was over 3; by 1910, the predicted ratio was only 1.35.

These findings indicate that wage-stretching was not a factor in the rise in wealth inequality in the late-nineteenth century. The gap between the wealth of the unskilled and the wealth of the skilled was shrinking rather than expanding during this period. This seems at odds with the data on wages during this period assembled by Williamson and Lindert.<sup>32</sup> These data indicate that the ratio of skilled to unskilled wages was growing at an annual rate of 0.3 percent for the period 1869-1899 and 1 percent for the period 1899-1909. The two sets of findings can,

---

where  $\phi_i(\cdot)$  and  $\Phi_i(\cdot)$  represent the density function and cumulative distribution function of the standard normal distribution.

<sup>32</sup> Williamson and Lindert, American Inequality, p. 218.

however, be reconciled. Wages are only one determinant of wealth. An individual's wealth holdings at any given point in time reflect the cumulative effects of inheritance, past and current labor earnings, savings rates, portfolio choice, and returns on assets. Wage inequality clearly has an impact on the distribution of wealth, but wage inequality and wealth inequality need not always exhibit the same trends and patterns. The wealth data for Massachusetts indicate that even if the wage gap between the unskilled and skilled was growing at the end of the nineteenth century, other factors were leading the wealth gap between these groups to fall.

While inequality did not increase between groups, the final column of Table 4 reveals that inequality did increase *within* groups. Figure 2 charts the group-level Theil measures for all the decompositions. Inequality within urban areas increased substantially, particularly in the antebellum period. Clearly, as the urban population grew, it became more diverse. The native-born population also experienced increasing inequality in the last decades of the nineteenth century. At the same time, wealth inequality within the foreign population declined. In 1850, the Theil measure for the foreign-born population was almost three times that for the native-born population. By 1910, the Theil measure for the foreign-born population was less than ten percent higher than that for the native-born population. The data on the inequality within different age groups also reveal interesting trends. The groups experiencing the greatest increases in inequality during the period were the 50 to 59 and 60 to 69 year-old age groups. Inequality was also increasing between 1870 and 1910 for all occupation groups except farmers. For the unskilled, the rise in inequality between 1870 and 1910 was a reversal of a decline in inequality between 1850 and 1870. Wealth inequality for the unskilled in 1910 was still

somewhat below that observed in 1850. But for skilled and white collar workers, wealth inequality in 1910 was substantially higher than in 1850.<sup>33</sup>

The finding of increasing inequality within groups remains even when considering more narrowly-defined population groups. The bottom of Table 4 presents decompositions in which population sub-groups are defined by pairs of characteristics such as urban/rural residence and nativity, and age-group and occupation. Even using these more narrowly defined sub-groups, we find that within-group inequality was increasing between 1870 and 1910.

The increases in within group inequality are an interesting phenomenon, which have not been anticipated by any of the major hypotheses for growing inequality during the nineteenth century. Since economists have considerable evidence indicating that labor markets were reasonably well integrated by the second half of the nineteenth century, within regions such as New England, it is reasonable to believe that competition would have approximately equalized wages within occupations.<sup>34</sup> If correct, one must look to beyond wages to other powerful forces that affected wealth ownership. Conceivably, unemployment rates could be a candidate that affected earnings and therefore wealth. Much of the rising inequality, however, occurred through growing concentration of wealth near the upper end of the wealth distribution--at levels of wealth ownership only modestly affected by labor market earnings. As shown in Table 3, between 1820 and 1910 the share of wealth held by the top 20 percent of the wealth distribution increased from 72 to 98 percent and the share held by the top 1 percent increased from 20 to 35

---

<sup>33</sup>We also performed decompositions using more disaggregated occupation categories: professional, business, other white collar, skilled, semi-skilled, and farmers. The qualitative results were the same: inequality within occupation groups was increasing between 1870 and 1910 even while inequality between groups was decreasing.

<sup>34</sup> Atack, Bateman and Margo, "Rising Wage" report an increase in the dispersion of unskilled wages between 1860 and 1880, but most of this is accounted for by the gap between north and south, as opposed to dispersion within the north.

percent. Variations in rates of returns on widely held assets such as land seem like an attractive hypothesis for study. Spectacularly high returns on some entrepreneurial activities, a central feature of the Gilded Age, also holds promise.

### **Concluding Remarks**

This paper assembles new data and new methods of analysis for studying wealth inequality trends in industrializing America. Records of household heads from the census matched with real and personal property tax records for Massachusetts reveal that the Gini coefficient and the Theil entropy measure of inequality were highly correlated. The Theil measure approximately doubled over the period from 1820 to 1910, a gain that was divided about evenly between the antebellum and the postbellum periods. A surge between 1870 and 1900 dominated the growth in inequality following the Civil War.

We test several popular hypotheses for inequality growth using decompositions of the Theil inequality measure and tobit regressions. The most successful hypothesis was put forward by Kuznets regarding composition effects. Population redistribution to urban areas alone would have led to a 53 percent increase in the Theil measure from 1820 to 1850 and a 22 percent increase from 1870 to 1910, holding inequality between and within urban and rural areas unchanged. Proposed labor market explanations, including sectoral shift (that led to higher wages in non-agricultural relative to agricultural sectors), biased technological change, and immigration are inconsistent with the fact that inequality between occupational groups was declining in the last decades of the century. Wealth accumulation patterns by age are also inconsistent with the hypothesis of child default on responsibilities for old age care, at least during the second half of the nineteenth century. To the extent that age patterns of wealth

ownership changed after 1850, relatively more wealth accumulated at very old ages as opposed to the 30s and 40s when parents would have amassed assets other than children in planning retirement.

None of the popular explanations of rising inequality in the nineteenth century can explain the rising inequality within groups that we find. Therefore, it is worth investing resources to develop an alternative hypothesis that may account for the salient facts. We are led to propose a new mechanism based on luck, rents, and entrepreneurship, which might explain the growing heterogeneity within occupations and within rural and urban areas. Any explanation must also incorporate the important role of real estate in total taxable wealth (about 70 percent), and the growing share this type of asset assumed in the portfolios of the very rich. A quote attributed to John Jacob Astor, made shortly before his death in 1848, is inspiring in this regard: “Could I begin life again, knowing what I now know, and had money to invest, I would buy every foot of land on the island of Manhattan.”<sup>35</sup> If a shrewd investor such as Astor was surprised by the vast returns to be made on real estate, surely large numbers of ordinary citizens were unable to judge as well. Many nineteenth-century Americans played the real estate game, which took on the dimensions of a repeated lottery whose outcomes were driven by unpredictable factors such as immigration and the precise routes taken by transportation systems. New entrepreneurial opportunities also appeared in America’s industrializing economy, with vast sums wagered, won, and lost in developing new industries such as railroads, oil, steel, and finance, which signify the Gilded Age.

Longitudinal data on wealth may be useful in testing at least some of the predictions of this hypothesis. Linked household data has already been assembled for the period 1850 to 1860

---

<sup>35</sup> Jackson, Encyclopedia, p. 63.

by Richard Steckel and by Joseph Ferrie.<sup>36</sup> These data were created from census schedules and include the data on wealth collected in those census years. Similar longitudinal data may be created for the late-nineteenth century using census schedules and property tax records. These data can be used to construct wealth transition matrices which describe the degree of wealth mobility in the two periods. If the hypothesis is correct, we should observe greater wealth mobility in the later period, and this greater mobility should remain even when controlling for occupation, size of place, age, and nativity.

---

<sup>36</sup> Steckel, "Census Matching," and Ferrie, Yankeys Now.



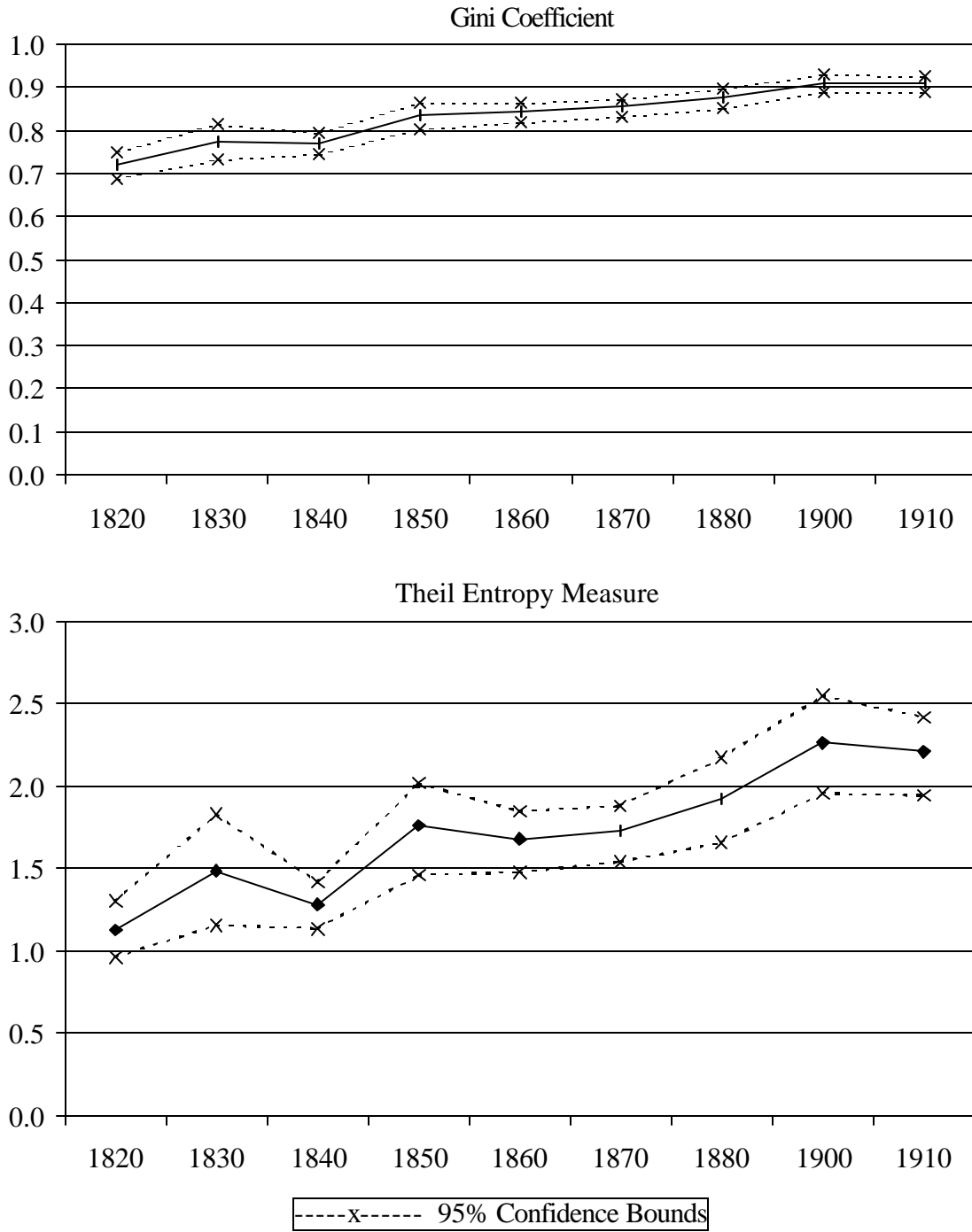
## REFERENCES

- Atack, Jeremy, Fred Bateman and Robert A. Margo, "Rising Wage Dispersion in American Manufacturing, 1860 to 1880." Unpublished Working Paper, Nashville, TN, January, 2000.
- Atack, Jeremy and Fred Bateman, "Egalitarianism, Inequality, and Age: The Rural North in 1860," Journal of Economic History 41, No. 1 (1981), 85-93.
- Bullock, Charles J., "The Taxation of Property and Income in Massachusetts," Quarterly Journal of Economics 31 (1916), 1-61.
- Conley, Timothy G. and David W. Galenson, "Quantile Regression Analysis of Censored Wealth Data." Historical Methods 27 (1994), 149-165.
- Commonwealth of Massachusetts, Aggregate of Polls, Property, Taxes, Etc., (various years).
- Cowell, Frank A., "The Structure of American Income Inequality." Review of Income and Wealth 30 (1984), 351-375.
- Cowell, Frank A. and Kiyoshi Kuga, "Additivity and the Entropy Concept: An Axiomatic Approach to Inequality Measurement." Journal of Economic Theory 25 (1981), 131-143.
- Ely, Richard T., Taxation in American States and Cities (New York: Thomas Y. Crowell and Co, 1888).
- Ferrie, Joseph P., Yankeys Now: Immigrants in the Antebellum U.S., 1840-1860 (New York: Oxford University Press, 1999).
- Foster, James E. "Inequality Measurement." Proceedings of Symposia in Applied Mathematics 33 (1985), 31-68.
- Jackson, Kenneth T., Encyclopedia of New York City (New Haven: Yale University Press, 1995).
- Jones, Alice Hanson, The Wealth of a Nation to Be (New York: Columbia University Press, 1980).
- Kuznets, Simon, "Economic Growth and Income Inequality," American Economic Review 45 (1955), 1-28.
- Kuznets, Simon, Shares of Upper Income Groups in Income and Savings (New York: National Bureau of Economic Research, 1953).
- Lampman, Robert J., The Share of Top Wealth-Holders in National Wealth, 1922-1956 (Princeton: Princeton University Press, 1962).

- Lindert, Peter H., "Three Centuries of Inequality in Britain and America." In A.B. Atkinson and F. Bourguignon (eds.), Handbook of Income Distribution, Volume I (Amsterdam: Elsevier Science B.V., 2000). Pp. 167-216.
- Main, Gloria, "Inequality in Early America: The Evidence of Probate Records from Massachusetts and Maryland." Paper presented at the Cliometrics Conference, Madison Wisconsin, April 1976.
- Main, Jackson Turner, The Social Structure of Revolutionary America (Princeton: Princeton University Press, 1965).
- Mills, Jeffrey A. and Sourushe Zandvakili, "Statistical Inference via Bootstrapping for Measures of Inequality." Journal of Applied Econometrics 12 (1997), 133-150.
- Nichols, Philip, Taxation in Massachusetts: A Treatise on the Assessment and Collection of Taxes, Excises, and Special Assessments under the Laws of the Commonwealth of Massachusetts (Boston: The Financial Publishing Co, 1913).
- Plehn, Carl C., Revenue Systems of State and Local Governments (Washington: GPO, 1907).
- Ransom, Roger L. and Sutch, Richard, "Two Strategies for a More Secure Old Age: Life-Cycle Saving by Late Nineteenth-Century American Workers." NBER Summer Institute on Development of the American Economy, Cambridge, July 1989.
- Soltow, Lee, Men and Wealth in the United States, 1850-1870 (New Haven : Yale University Press, 1975).
- Soltow, Lee, Distribution of Wealth and Income in the United States in 1798 (Pittsburgh: University of Pittsburgh Press, 1989).
- Soltow, Lee, "Inequalities in the Standard of Living in the United States." In Robert E. Gallman and John J. Wallis (eds.), American Economic Growth and Standard of Living before the Civil War (Chicago: University of Chicago Press, 1992). Pp. 265-308.
- Steckel, Richard H., "Census Matching and Migration: A Research Strategy," Historical Methods (Spring 1988): 52-60.
- Steckel, Richard H., "Census Manuscript Schedules Matched with Property Tax Lists: A Source of Information on Long-Term Trends in Wealth Inequality," Historical Methods 27 No. 2 (Spring 1994), 71-85.
- Street, Alfred Billings, A Digest of Taxation in the States (Albany: Weed, Parsons, 1863).
- Sutch, Richard, "All Things Reconsidered: The Life-Cycle Perspective and the Third Task of Economic History," Journal of Economic History 51 No. 2 (June 1991), 271-88.
- Williamson, Jeffrey G. and Lindert, Peter H., American Inequality: A Macroeconomic History (New York: Academic Press, 1980).

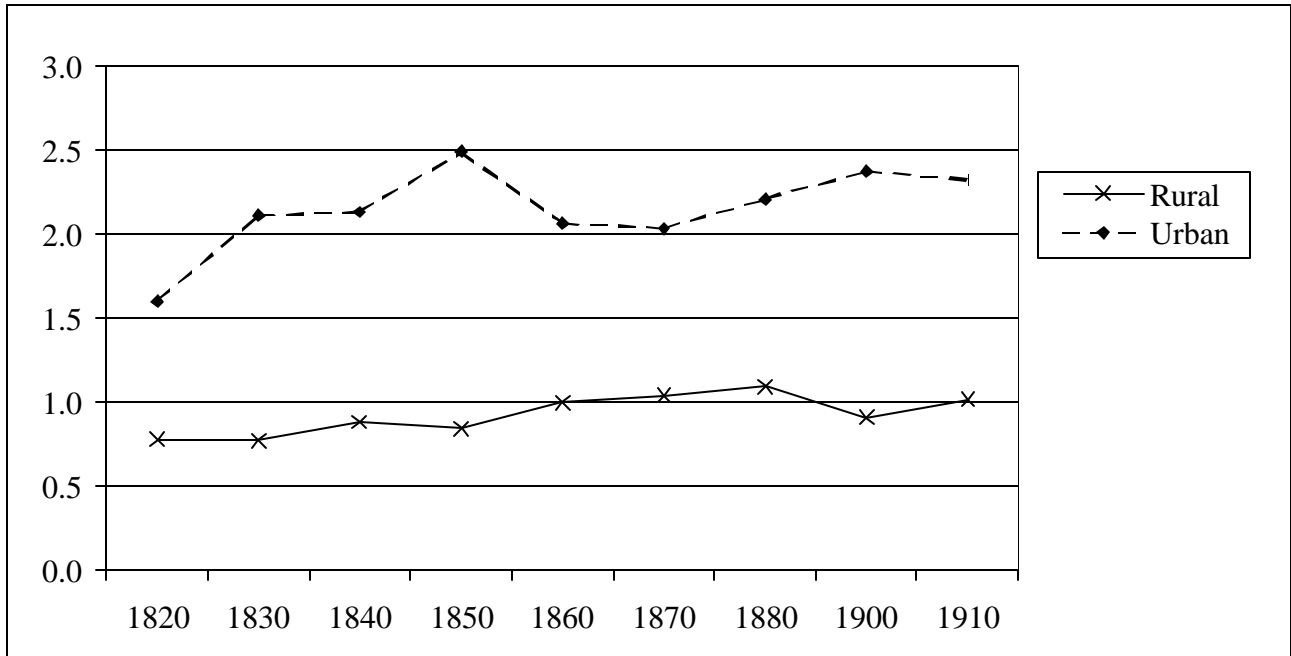
Wright, Carroll Davidson, The History and Growth of the United States Census (Washington: U.S. Government Printing Office, 1900).

**Figure 1.—Wealth Inequality, Massachusetts 1820-1910**



**Figure 2.—Theil Entropy Measures for Population Sub-groups**

Urban/Rural



Nativity

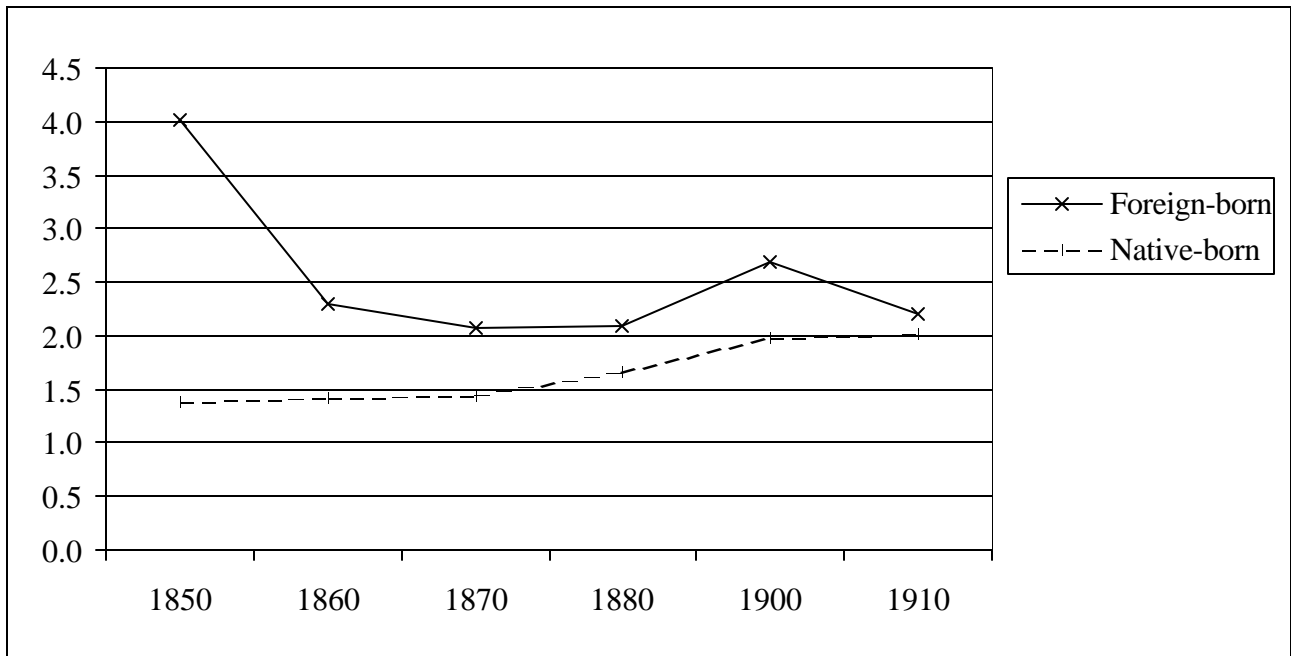
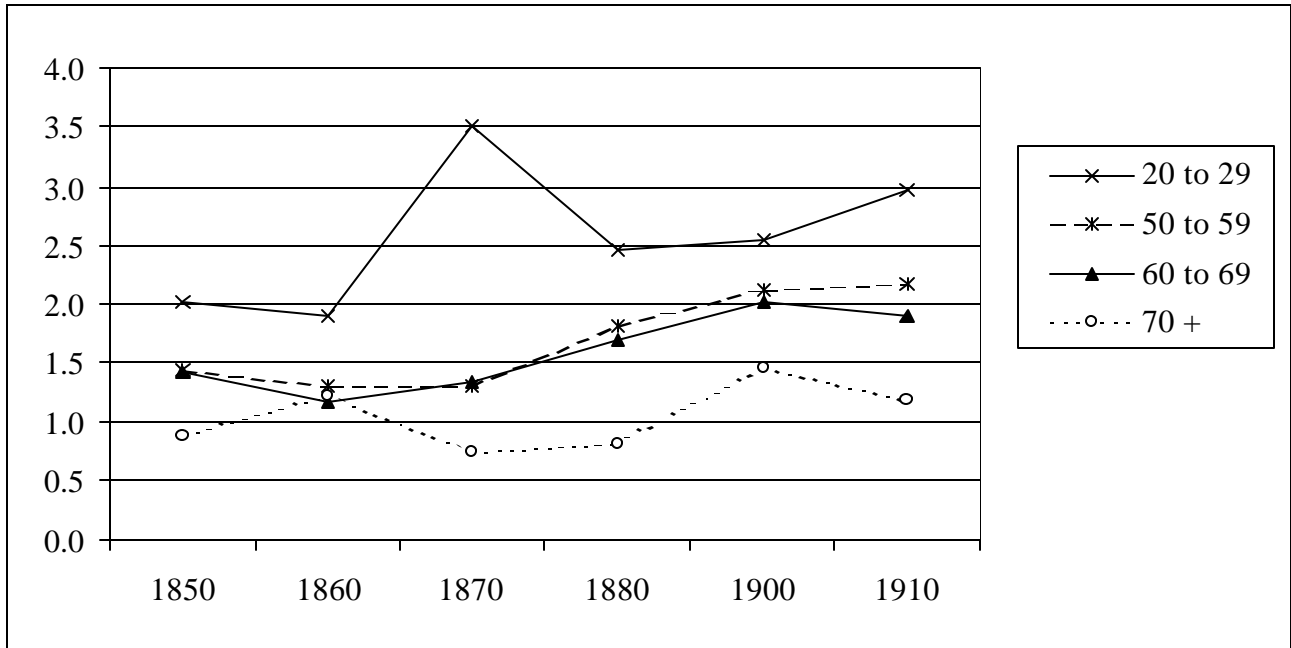
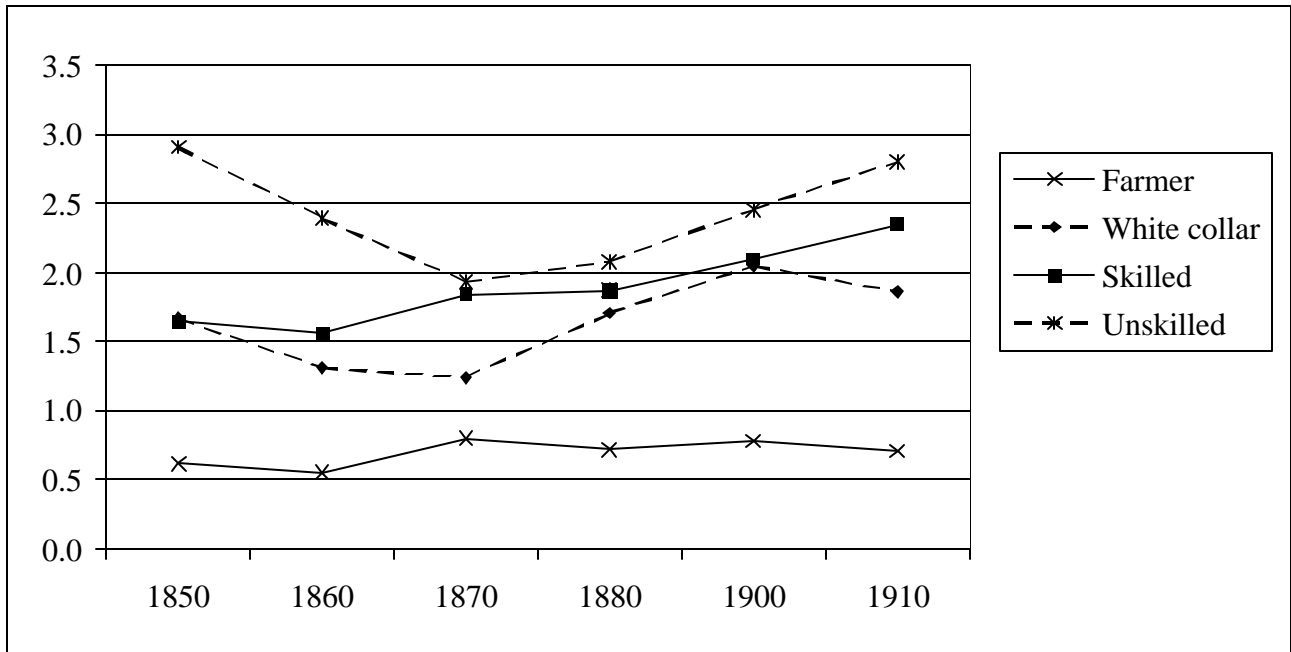


Figure 2.—Continued

Age Groups



Occupation



**Table 1.—Descriptive Statistics**

	1820	1830	1840	1850	1860	1870	1880	1900	1910
Urban	22.8%	31.1%	37.9%	50.6%	59.5%	66.7%	74.7%	86.0%	89.0%
Age of head									
< 20				0.2	0.1	0.3	0.3	0.7	0.1
20 - 29				14.9	14.7	13.9	11.5	12.9	13.5
30 - 39				32.2	32.4	27.2	27.0	29.1	29.4
40 - 49				25.8	23.3	25.0	25.8	22.0	25.8
50 - 59				14.3	16.2	19.8	17.2	17.3	17.8
60 - 69				8.5	9.4	9.2	11.9	12.0	8.4
70 +				4.0	3.9	4.6	6.2	6.1	5.0
Black				0.8	1.1	1.1	1.2	2.1	2.1
Foreign-born				24.5	29.8	39.6	42.1	45.6	42.8
Illiterate				4.8	5.3	7.7	8.1	5.6	6.1
Occupation of head									
Farmers				23.0	16.1	11.6	9.3	8.5	4.9
White collar				15.4	17.5	18.9	21.3	29.4	32.5
Skilled				26.9	33.3	28.2	27.4	22.9	23.0
Unskilled				29.7	29.6	37.5	34.4	32.9	32.3
Other				5.0	3.6	3.8	7.6	6.3	7.4
No match in tax records	34.2	34.1	37.7	51.2	54.7	58.2	66.2	70.1	73.6
No. of Observations	1016	989	977	1023	1005	1017	1020	977	1003

Note: Rural and urban observations were weighted by the rural and urban population shares in the state.

**Table 2.—Share of Real Estate Wealth in Total Taxable Wealth, Male Household Heads, Massachusetts, 1820-1910**

	Sample Aggregate	Average real estate share for individuals in:		
		Top 20%	Top 5%	Top 1%
1820	76.9%	79.1%	62.7%	36.8%
1830	66.5	75.1	58.6	32.1
1840	75.4	78.8	75.8	55.3
1850	64.6	75.8	72.1	56.5
1860	55.5	71.7	50.8	29.0
1870	51.7	62.3	48.2	35.5
1880	73.3	82.7	66.2	62.9
1900	71.9	78.4	76.7	64.9
1910	68.7	74.8	71.3	70.0

**Table 3.—Distribution of Total Taxable Wealth, Male Household Heads,  
Massachusetts, 1820-1910**

	N	Share of Wealth held by:			Gini	Theil Entropy
		Top 20%	Top 5%	Top 1%		
1820	1016	72.0% (1.7%)	40.5% (2.7%)	20.3% (2.8%)	0.720 (0.015)	1.125 (0.086)
1830	989	77.6 (2.2)	49.2 (4.5)	28.9 (5.3)	0.775 (0.020)	1.486 (0.175)
1840	977	78.3 (1.6)	45.0 (2.8)	20.0 (2.4)	0.771 (0.013)	1.282 (0.072)
1850	1023	85.8 (1.7)	55.7 (4.5)	33.4 (4.6)	0.836 (0.016)	1.761 (0.147)
1860	1005	88.1 (1.4)	55.7 (3.5)	27.0 (3.2)	0.844 (0.012)	1.679 (0.096)
1870	1017	90.1 (1.2)	56.7 (3.3)	27.2 (2.9)	0.856 (0.011)	1.730 (0.086)
1880	1020	93.7 (1.2)	60.3 (4.1)	29.1 (5.0)	0.877 (0.012)	1.924 (0.136)
1900	977	97.3 (0.9)	70.5 (4.0)	37.2 (5.6)	0.911 (0.011)	2.264 (0.157)
1910	1003	98.3 (0.8)	68.7 (3.8)	35.0 (4.6)	0.910 (0.010)	2.207 (0.124)

Note: Numbers in parentheses are approximate standard errors obtained by bootstrapping.



**Table 4.—Decompositions of Changes in Theil Entropy Measure**

Sub-group partition	Total percent change in T	Percent change in T accounted for by changes in:		
		Population shares	Sub-group mean wealth	Within-group Inequality
<u>1820-1850</u>	56.5%			
Urban/rural		53.4%	-29.2%	32.3%
<u>1870-1910</u>	27.6			
Urban/rural		21.9	-5.5	11.2
Nativity		-5.5	2.5	30.6
Age		-5.5	12.4	20.8
Occupation		41.4	-40.0	26.1
Urban/Rural &: Nativity		25.0	-14.6	17.2
Age		29.2	-9.5	7.9
Occupation		54.7	-47.5	20.4
Nativity &: Age		-22.3	29.4	20.5
Occupation		31.6	-28.6	24.6
Age & Occupation		38.8	-31.3	20.1

**Table 5.—Tobit Regression Results for Log Wealth, 1850, 1870 and 1910**

	<u>1850</u>		<u>1870</u>		<u>1910</u>	
	Coefficients	Predicted ratio of group wealth to base wealth	Coefficients	Predicted ratio of group wealth to base wealth	Coefficients	Predicted ratio of group wealth to base wealth
Age of head						
< 20	-33.377 (22085)	0.164	-36.460 (23465)	0.390	0.333 (11.645)	1.041
30 - 39	0.865 (0.679)	1.514	3.784 (1.149)	3.730	4.219 (1.831)	2.110
40 - 49	3.145 (0.686)	5.872	5.678 (1.147)	9.922	7.734 (1.832)	6.443
50 - 59	2.812 (0.767)	4.708	6.992 (1.173)	22.223	7.078 (1.898)	5.026
60 - 69	2.396 (0.881)	3.611	7.352 (1.344)	28.213	8.087 (2.149)	7.427
70 +	3.581 (1.136)	7.932	7.641 (1.631)	34.345	12.106 (2.423)	56.412
Occ. of head						
Farmers	3.980 (0.660)	10.554	4.719 (1.011)	5.880	8.748 (1.892)	9.838
White collar	5.175 (0.683)	26.278	4.982 (0.866)	6.749	7.666 (1.205)	6.273
Skilled	2.200 (0.608)	3.199	1.498 (0.795)	1.529	2.094 (1.309)	1.353
Other	1.564 (1.073)	2.204	3.606 (1.534)	3.440	4.340 (1.828)	2.174

**Table 5.—Continued.**

	<u>1850</u>		<u>1870</u>		<u>1910</u>	
	Coefficients	Predicted ratio of group wealth to base wealth	Coefficients	Predicted ratio of group wealth to base wealth	Coefficients	Predicted ratio of group wealth to base wealth
Urban	-2.872 (0.493)	0.370	-2.840 (0.708)	0.593	-5.447 (1.331)	0.687
Black	-0.995 (2.490)	0.664	-5.437 (3.908)	0.466	-2.809 (3.911)	0.779
Foreign-born	-4.957 (0.698)	0.246	-2.701 (0.714)	0.604	-0.313 (0.941)	0.965
Illiterate	-2.968 (1.747)	0.361	-0.477 (1.287)	0.894	-2.809 (2.172)	0.779
Intercept	-0.710 (0.753)		-4.460 (1.229)		-10.001 (2.214)	
Scale ( $\sigma$ )	5.372 (0.197)		6.471 (0.305)		8.423 (0.512)	
No. of Obs.			1017		1003	

Notes: Dependent variable is  $\ln(\text{wealth} + \$1)$ . The expected observed wealth of different groups used to construct the predicted wealth ratios were calculated using the following formula:

$$E[\ln(w_i + 1)] = \Phi_i(x_i' \beta / \sigma) x_i' \beta + \sigma \phi_i(x_i' \beta / \sigma)$$

where  $\phi_i(\cdot)$  and  $\Phi_i(\cdot)$  represent the density function and cumulative distribution function of the standard normal distribution.