Rochester Center for
Economic Research

The Rise and Fall of Unions in the United States

Emin Dinlersoz and Jeremy Greenwood

Working Paper No. 596

August 2016

UNIVERSITY OF
ROCHESTER

The Rise and Fall of Unions in the United States

Emin Dinlersoz[a,*], Jeremy Greenwood[b,†]

[a] U.S. Census Bureau; [b] University of Pennsylvania

August 2016

Forthcoming in Journal of Monetary Economics

Abstract

Union membership in the United States displayed a ∩-shaped pattern over the 20th century, while income inequality sketched a ∪. A model of unions is developed to analyze these facts. There is a distribution of productivity across firms in the economy. Firms hire capital, plus skilled and unskilled labor. Unionization is a costly process. A union chooses how many firms to organize and the union wage. Simulation of the model establishes that skill-biased technological change, which affects the productivity of skilled labor relative to unskilled labor, can potentially explain the observed paths for union membership and income inequality.
Keywords: Mass Production, Computer Age, Skill-Biased Technological Change, Income Inequality, Union Membership
JEL classification: J51, J24, L23, L11, L16, O14, O33

1. Introduction

In 1910, around 10% of the American workforce in the non-agricultural private sector were union members. As shown in Figure 1, the percentage of union members rose until the middle of the century, reaching its apex at roughly 40%. It then began a slow decline. By

*Corresponding author: Emin Dinlersoz, Center for Economic Studies, U.S. Census Bureau, 4600 Silver Hill Rd., Suitland, MD 20647. Tel: (301) 763-7889. E-mail: emin.m.dinlersoz@census.gov

†Any opinions and conclusions expressed herein are those of the authors and do not necessarily represent the views of the U.S. Census Bureau. All results have been reviewed to ensure that no confidential information is disclosed. We thank Ricardo Reis and an anonymous referee for detailed reviews that improved the article. Omer Acikgoz, Henry Hyatt, Kristin McCue, Lee Ohanian, James Schmitz, and Stijn Vanormelingen provided useful comments. William Lefevre of Walter P. Reuther Library of Labor and Urban Affairs kindly supplied the membership figures for the UAW union.

the end of the century, only about 8% of American workers belonged to a union. Income inequality followed a different path. At the beginning of the 20th century, the income share of the top 10% was 40%. This measure of income inequality first declined, hitting a low of 31% around mid-century. It then steadily increased to about 42% around 2000.[1] What could have caused the ∩-shaped pattern of union membership and the ∪-shaped one for income inequality? Are they related?

The hypothesis advanced here is that skill-biased technological change underlies the rise and fall in union membership, along with the decline, and then the surge, in income inequality. The beginning of the 20th century witnessed a shift away from an artisan economy toward one transformed by mass production. This transformation favored unskilled labor. The premium for skill declined.[2] Unskilled labor is homogenous almost by definition, making it easier to unionize than skilled labor. When the demand for unskilled labor rises, there is a larger payoff to unionizing it. These trends started to shift at the midpoint of the century. The Second Industrial Revolution was petering out and the Information Age was dawning. Transistors and silicon chips meant that automatons could replace the hoards of unskilled workers laboring on factory and office floors. These developments represented a reversal of the patterns observed earlier in the 20th century.

To explore the connection between technology, unions and income inequality, a general equilibrium model of unionization is developed. A single union makes two interconnected decisions. First, it picks a common wage rate for its members. Second, the union selects which firms to organize. Unionization is a costly process. Firms sell output in a competitive market. They hire both skilled and unskilled labor. These inputs are substitutable to some extent. When the productivity of unskilled labor is high (relative to skilled labor) the union

[1] All supplementary material is available as an Online Appendix on Science Direct. The income inequality measure is before individual income taxes–see the Online Appendix for more detail. Therefore, changes in the progressivity of income taxation do *not* account for the ∪-shaped pattern in income inequality. The rise in inequality since the 1970s is well documented and holds for a wide variety of inequality measures–see Juhn, Murphy, and Pierce (1993) for an early documentation of this trend for many measures of wage inequality. Some other time series measures of income inequality are shown in Figure 4. They all display the same ∪-shaped pattern.

[2] Interestingly, Goldin and Katz (2008, Figure 8.1, p. 290) report a ∪-shaped pattern for the college-graduate wage premium for the period of study here–see Figure 4. Somewhat surprisingly, they also show that during the first part of the twentieth century the high-school graduate wage premium actually fell; i.e., the return to a less-than-high-school education rose. These facts fit well into the framework laid out here.

can pick a high wage. It also pays to organize more firms. Firms differ in productivity, so when organizing labor the union selects the most profitable firms. Those firms that are not unionized can hire labor in a competitive market.

The modeling of unions builds upon the work of MacDonald and Robinson (1992). They present a model of the extent of unionization in a competitive industry where all firms are the same. The key ingredients of their model are: (i) unionization is a costly activity; (ii) unions must offer their members a wage net of dues that exceeds the competitive one; (iii) the union wage must allow organized firms to make non-negative profits. MacDonald and Robinson (1992) model an industry in partial equilibrium, and start off at the level of a firm's cost function. Modeling skill-biased technological change requires, instead, starting off from a firm's production function that uses both skilled and unskilled labor. In addition, analyzing the implications of this form of technological change for the income distribution requires a general equilibrium approach that embeds unions, as well as heterogeneous individuals and firms. These elements are needed to model the rise and fall of unions.

The hypotheses proposed here is supported by historical evidence regarding the evolution of unionization and skill-biased technological change. This evidence is laid out in detail with particular attention to the transformation of the U.S. economy over the 20th century initially by mass production, and later, by computerization and automation. The developed model is also calibrated and simulated to see whether it is capable of explaining the time paths of unionization and income inequality. It is. The implied pattern of skill-biased technological progress is in line with the qualitative picture painted by the historical evidence.

Acemoglu, Aghion and Violante (2002) also analyze how skill-biased technological change can lead to deunionization. Their framework is very different from the one developed here. In particular, there are two sectors in the economy, one unionized, the other non-unionized. Skilled workers only work in the non-unionized sector. Unskilled labor can work in either sector. As the productivity of skilled workers relative to unskilled workers rises more people choose to become skilled and hence are employed in the non-unionized part of the economy. Their analysis is entirely theoretical in nature. Acikgoz and Kaymak (2014) embed a model of unionization into a Mortensen-Pissarides style job matching model. In their framework, workers differ both by their ability and skill levels. Firms observe both attributes, while unions see only the latter. They argue that a rise in the skill premium, which rewards both ability and skill, reduces the incentive for a skilled worker to join a union. The rise in the skill premium is also associated with unskilled workers becoming less productive. Firms find such workers less attractive to hire at high union wages. The current analysis stresses, by contrast, the interplay between firms and unions. Unions organize the most profitable slice of the firm productivity distribution. The size of this slice depends on the state of the production technology in the economy and the cost of union organizing. Furthermore, not only the fall, but also the rise, of unions is explained by technological change.

2. Technology, Unions, and Income Inequality

In a nutshell, the hypothesis advanced here is this: The Second Industrial Revolution breathed life to unions and, then, the Information Age took it away. To set the stage for the model, detailed historical evidence is now provided on the connection between technology, unionization, and income inequality in the United States. Specifically, it is argued that the rapid rise of unions during the first half of the 20th century was closely related to the deskilling of the workforce induced by the diffusion of mass production. Conversely, the decline of mass production, and the rise of automation and computerization, were instrumental in the rise of skilled labor and the fall of unions during the second half of the century. The patterns of unionization and income inequality over the course of the 20th century were both driven by a common factor: skill-biased technological change.

2.1. The Rise of Unions and the Decline in Income Inequality, 1913-1955

Mass production and Fordism were interchangeable terms at one time. In 1913 Ford's Highland Park plant became the first automobile factory to have a moving assembly line. It signalled the death of the craft production methods that characterized the previous century. The use of standardized parts, pioneered in the 19th century arms industry, eliminated time spent fitting inexact parts. The moving assembly line was inspired by the flow production techniques used in flour milling and meat packing. It reduced the unnecessary handling of the product associated with ferrying the work between production operations. The result was a greater specialization of labor.

At the beginning of the 20th century, automotive, carriage and wagon, and machine and metal-working workshops were artisanal in character. They had three types of workers: skilled mechanics, specialists, and laborers. The skilled mechanics undertook the productive operations and supervised the other workers. A census report stated that the "machinist, in its highest application, means a skilled worker who thoroughly understands the use of metal-working machinery, as well as fitting and working at the bench with other tools." Laborers were unskilled and did "manual labor that requires little or no experience or no judgement, such as shovelers, loaders, carriers, and general laborers." The semi-skilled specialist lay between these two categories. The census referred to them as "machinists, of inferior skill." It stated that "those who are able to run only a single machine or perhaps do a little bench work, are classified as second class machinists and grouped with machine tenders and machine hands." Meyer (1981, p. 13-14) describes how Ford engines were put together just before the assembly line was born:

> At the assembly bench, the skilled worker occupied a central place. He began with a bare motor block, utilized a wide range of mental and manual skills, and attached part after part. Not only did he assemble parts, but he also 'fitted' them. If two parts did not go together, he placed them in his vice and filed them to fit. The work routines contained variations in tasks and required considerable amounts of skill and judgment. Additionally, unskilled truckers served the skilled assemblers. When an assembler completed his engine, a trucker carried it away and provided a new motor block. The laborer also kept the assembler supplied with an adequate number of parts and components. Here, the division of labor was relatively primitive–essentially, the skilled and unskilled. Under normal conditions, a Ford motor assembler needed almost a full day of work to complete a single engine.

Mass production involved breaking down the manufacturing process into a series of elementary tasks and the transfer of skill to machines. Frederick W. Taylor wrote in 1903 that "no more should a mechanic be allowed to do the work for which a trained laborer can be used" and that "a man with only the intelligence of an average laborer can be taught the most difficult and arduous work if it is repeated; and this lower mental caliber renders him more fit than the mechanic to stand the monotony of repetition." A 1912 report of the American Society of Mechanical Engineers stated that "after the traditional skill of a trade, or the peculiar skill of a designer or inventor, has been transferred to a machine, an operator with little or no previously acquired skill can learn to handle it and turn off the product."

An 1891 sample of metal-working establishments in Detroit shows the importance of skilled labor in artisanal production. As Table 1 illustrates, mechanics accounted for nearly 40% of the workforce. Meyer (1981) feels that this pattern would have been characteristic of the early Ford Motor Company as well. The composition of the workforce at the Ford Motor Company had changed by 1913, as also shown in Table 1. Operators made up the majority of workers. These were unskilled specialists performing routine machine operations. Mechanics accounted for only a small portion of the workforce. The deskilling of the workforce is nicely related by Wolmack, Jones and Roos (1990, p. 31):

> The assembler on Ford's mass production line had only one task–to put two nuts on two bolts or perhaps attach one wheel to each car. He didn't order parts, procure his tools, repair his equipment, inspect for quality, or even understand what the workers on either side of him were doing. Rather, he kept his head down and thought about other things. The fact that he might not even speak the same language as his fellow assemblers or the foreman was irrelevant for the success of Ford's system.

Mass production was by no means limited to automobile manufacturing. During the 1910s and 1920s, mass production and assembly line techniques spread to several manufacturing industries, such as steel, tires, petrochemicals, plastics, aircraft, textiles, and cigarettes. They also transcended the boundaries of manufacturing into other industries, even those as diverse as film making. Storper (1989) provides a detailed account of how the film industry welcomed mass production practices during the 1920s as it moved from New York to California. During this period, the film industry transformed from one that closely replicated the craft methods of the theater into one that utilized mass production, supported by the creation of a large market and a highly standardized product; Hampton (1970) notes that during this period some films were in fact sold by the foot rather than on the basis of content. The transition from craft methods to mass production, however, was complex, particularly for labor. Meyer (1988) emphasizes the effects of this transition in the Allis-Chalmers Manufacturing Co., a major manufacturer of industrial equipment. The transition meant very different changes for management versus workers. For managers, the new production technologies implied a reduction in skilled labor, an increase in managerial involvement in production, and a decline in production costs. Workers, on the other hand, faced a loss of skills, a higher intensity and faster pace of work, and greater job insecurity because of changing production practices.

The reorganization of production processes in favor of unskilled labor implied that the patterns of unionization during the first half of the 20th century were closely tied to the diffusion of mass production. As opposed to the craft unions that prevailed earlier, the new trend was more in the form of industrial unionization, made possible by the masses of workers in big industrial plants. The demand for collective bargaining was driven by the deskilling of the workforce and the special environment created by mass production: repetitive work performed by relatively homogeneous workers with an ever increasing need for speed to expand production–the “speedup.” Edsforth and Asher (1995) argue that this increasingly demanding environment led a critical mass of workers to form the United Automobile Workers (UAW) union in 1936. The UAW union began collective bargaining in General Motors and Chrysler in 1937, and Ford in 1941–these dates come after the early diffusion of mass production technologies in the U.S. economy during the 1910s and 1920s.

Figure 1 plots the evolution of union membership in manufacturing. The rapid rise in union membership in the earlier part of the 20th century and the subsequent long decline resemble those in the entire private sector. Figure 1 also displays the rise and fall of membership for the UAW.[3] Note the steep rise in the number of union members shortly after the UAW was founded in 1936. Similar time paths were observed for union membership

[3] The figures pertain to dues-paying members.

in other industries. The United Rubber Workers (URW) union was formed in 1935 and experienced sharp gains in membership, before declining and becoming part of the United Steel Workers (USW) union in the second half of the century. The USW union signed its first contract with Carnegie-Illinois Steel in 1937 for a $5-a-day wage and benefits. As in the case of the automobile industry, the steel and rubber industries experienced major changes in the organization of production as they adopted mass production. The Bessemer and open hearth processes in steel production, and the Banbury mixer in tire production, paved the way for the adoption of mass production practices.[4] These inventions predate the rapid rise in union membership starting in the 1930s.[5] In particular, Nelson (1987, p. 48) points out that mass production in the rubber industry diffused so fast that by 1920 it was too late for laggards to catch up and maintain sufficient profits with respect to more advanced rivals that adopted it earlier.

The experience of the automobile, tire, and steel industries suggest that the diffusion of mass production started before the acceleration of unionization in these industries. Comprehensive data on the diffusion of mass production methods for the economy at large is not available, because there is no exact definition of what qualifies a business as a unit of mass production.[6] However, the time path of some industry aggregates are informative on the timing of the diffusion. One aspect of mass production is that it allowed for a rapid expansion of output and employment. Therefore, large shifts in employment and output should be observed in the first half of the 20th century for those industries typically thought of as adopters of mass production methods. Figure 2 displays the physical output for cars, raw steel, tires, and cigarettes. The output in each case increased sharply between 1910 and 1930, and continued to do so after a temporary fall due to the onset of the Great Depression.

[4] See Kahan (2014, p. 12) for a discussion of how the Bessemer process helped increase the rate of steel production and how it changed the worker's routines and required more of them. See Nelson (1987) for the adoption of mass production methods in the U.S. tire industry and its consequences for the organization of production.

[5] The Banbury mixer was invented in 1916 by Fernley H. Banbury. The Bessermer process was introduced to U.S. steelmaking during the 1880s. The open hearth process, also called Siemens-Martin process, was first invented in the 1860s and was utilized in U.S. steel production starting in the early 1900s.

[6] On this point, Henry Ford (1926) defines mass production as "the focusing upon a manufacturing project of the principles of power, accuracy, economy, system, continuity, speed and repetition." He emphasizes that "Mass production is not merely quantity production... Nor is it merely machine production ..."

The expansion of output in these industries preceded the rise in union membership in the 1930s. Many industries also experienced a growth in average plant size. The average output and employment of a plant in the tire industry grew nearly 7-fold and 4-fold, respectively, between 1921 and 1937, as shown in Figure 2. Figure 2 also highlights the dramatic rise in employment and output in the Ford Motor Company after the introduction of the assembly line in 1913.

As another proxy for the diffusion of mass production, Figure 3 plots the frequency of use for the phrase "mass production" in the print media, using the Google Books Ngram Viewer.[7] This frequency serves as one possible proxy for the importance of mass production as a social and industrial phenomenon. One would expect that as the underlying technologies spread the phrase was more frequently referred to in the print media (see also Figure A1 in the Online Appendix).[8] It indeed started to gain popularity in mid-1910s, around the time Ford's Highland Park plant introduced assembly lines. In contrast, union membership did not start its surge until the 1930s, with the exception of a short-term rise and fall due to the escalated demand for the products of some unionized industries, such as coal and steel, during the years of World War I. Note also that the popularity of "mass production" peaked before union membership did.[9]

The rise of mass production and unionization was accompanied by a sharp fall in income inequality–see Figure 1. This fall is also apparent if one considers alternative ways of looking at inequality. Figure 4 presents the evolution of the income share of the top 1 and 5%, which are very similar to that of the top 10% in Figure 1, regardless of whether income includes

[7] The Google Ngram Viewer is a big data application designed to perform searches over a large sample of books written in English and published in the United States. It can be used to provide the relative frequency of the use of a given phrase among all phrases of the same word count occurring in all the books searched, normalized by the book counts in any given year to account for the growing number of books published over time. The details of this search tool are available at https://books.google.com/ngrams/info (Last accessed: August, 2015). For some examples of the increasing use of Google Ngram Viewer in scientific research, see Finke and McClure (2015), Michel et al. (2011), and Zheng and Greenfield (2015).

[8] Figure A1 provides an example of how the actual time series for a given quantity coincides with the time series for the frequency of its Google Books Ngram proxy. The frequency of the term "labor unions" over time follows actual union membership rate very closely. The correlation between the two series is quite high, 0.92.

[9] The pattern in Figure 3 is very similar if one uses the term "assembly line" in conjunction with "mass production."

capital gains or not. The share in lower and middle income ranges (1st-60th percentiles) also rose during the period 1929-1960. Similarly, the inverted Pareto-Lorenz coefficient, another measure of inequality, followed a ∪-shape. Note also that the 90-10 log-wage gap, and the returns to a year of high school and college, fell sharply between 1920 and 1950. The shrinking wage gap is consistent with a rise in the relative productivity of lower-wage, unskilled labor. The decline in the returns to schooling is also consistent with a fall in the relative demand for skilled workers, which may have reduced the incentives for education. The patterns in income inequality are also robust when other proxies are considered based on the popularity of a variety of terms describing individuals' income levels.[10] Finally, Figure 5 also offers a picture consistent with the hypothesis advanced here. It contains several measures aimed at capturing various dimensions of skill-biased technological change. During the first half of the 20th century, both the number of unskilled workers relative to skilled workers, and the number of production workers relative to all employees in manufacturing were high. The former actually increased between 1920 and 1940, and the latter between 1930 and 1950.

2.2. The Fall of Unions and the Rise in Income Inequality, 1955-

In 1952 MIT publicly demonstrated an automatic milling machine. The machine read instructions from a paper punch tape. The instructions were fed to servo-motors guiding the position of the cutting head of the machine relative to the part being manufactured along the x, y and z axes. Feedback from sensors regulated the process. By changing the instructions the machine could manufacture a different part. Such a "flexible machine" could make small batches of many different parts. The world had entered the age of numerically-controlled machines. Numerically-controlled machines were slow to catch on. Programming an early form of such machines was a time consuming task. Standardized languages were developed

[10] Figure A2 in the Online Appendix plots the frequency of the use of the term "rich people" relative to the terms "rich people," "middle class" and "poor people" combined. The initial fall in the relative frequency of "rich people" during the period 1930-1950 and the rise starting in the 1970s are consistent with the patterns in Figures 1 and 4. Other proxies in Figure A2 exhibit similar time paths. As in the case of Figure A1, the conformity between Figures 1, 4 and A2 further suggest that Google Ngram-based proxies are capable of capturing the salient features of the time series for the underlying variable of interest.

for programming automated machine tools by the 1960s. At the same time, the arrival of less expensive computers in the 1960s made them economical. The separation of software from hardware also lowered the costs of implementing numerical control systems. As calculating power increased, computers could aid the design of products. Computers could also be used for planning and managing business in addition to running the machines on the factory floor. In fact, sometimes they could automate virtually the entire business. The use of computers reduced the need for unskilled labor in factories and offices.

Mass production is an inflexible system. It is difficult to change a product or the manufacturing process once an assembly line has been instituted. As Henry Ford said "Any customer can have a car painted any color that he wants so long as it is black." This didn't suit Japanese manufacturing in the early postwar period, which had small production runs. The dies used in presses to shape metal parts had to be switched frequently. It took specialists in an American plant a day to change dies. Dies weighed tons and had to be set in the presses with absolute precision. Otherwise, defects would appear in the manufactured parts. In the 1940s and 50s, Taiichi Ohno, Toyota's chief production engineer, perfected a simple system where they could be changed in minutes. Since the presses had to remain idle while the dies where changed, Ohno reasoned that the production workers could do this. Furthermore, they could check the manufactured parts for defects thereby catching mistakes early on in the production process. Quality control was at the end of the process in the typical mass production facility. Over time, Toyota's production system gradually evolved to one where teams of workers were responsible for segments of the assembly line. Besides production, they looked after housekeeping, minor machine repairs and quality checking for their section of the line. According to Wolmack, Jones and Roos (1990), in a mass production automobile plant about 20% of the area and 25% of the working time are devoted to fixing mistakes. This is eliminated in a Toyota "lean production" facility. The Toyota production system favors skilled workers rather than unskilled ones. It has now been widely adopted in manufacturing.

The upshot of computerization in production and new organizational structures was that the demand for unskilled labor fell relative to the demand for skilled labor. The shift away from mass production had an adverse effect on those unions comprised of unskilled workers. Piore (1986) highlights this effect and argues that the decline of mass production as a form of technology and a means for industrial prosperity for unskilled workers had dire consequences for unions. As the economy moved from large scale, standardized production with abundant use of unskilled labor to smaller batches of production with a higher skill content, union organizing opportunities dwindled. Consistent with these trends, the proportion of unskilled and production workers in the U.S. economy, and their hours worked, all declined starting in the second half of the 20th century, as illustrated in Figure 5. In addition, Figure 3 indicates that the popularity of the phrase "mass production" peaked and then started to wane before union membership did. This timing suggests that the fall of mass production likely led that of unions. Figure 3 also includes measures of the popularity of the terms describing technologies that favor skilled labor, such as computers, automation, and robots. The frequency of use for these terms in the print media started to take off during the early 1950s, with an acceleration throughout the 1980s.[11] While union membership dropped somewhat during the mid-1940s (driven in part by the ending of World War II), it rebounded quickly, and its fast decline did not commence until the 1970s. The particularly steep decline in union membership during the 1980s also coincides with the sharp increase in the frequency of the terms describing advanced technology.

As technology continued to favor skilled workers during the second half of the 20th century, income inequality increased–see Figures 1 and 4, and Figure A2 in the Appendix. The acceleration in skill-biased technological change can be seen in the measures plotted in Figure 5. The relative price of equipment and software fell sharply, and the share of information processing equipment in private investment rose. These trends indicate a rise in the adoption of technologies that complement skilled labor. The relative price series is often used as a measure of skill-biased technological progress–see Krusell et al. (2000) and Cummins and Violante (2006). At the same time, the number of skilled workers, and their hours of work relative to those of unskilled, both rose, and the relative number of production

[11] It makes little difference to the overall picture if one adds other phrases describing technology, such as "computer-aided manufacturing," "flexible manufacturing," "computer-aided design," "numerically-controlled machines," inter alia.

workers fell.

2.3. Which Came First: Technological Change or Unionization?

The previous sections laid out the historical facts in support of the role of technology in driving, first, the unprecedented rise of unions and then, the fall. An alternative hypothesis is that the rise and fall of unions dictated the direction of technological change. Under this alternative, forces other than technology, such as the union-friendly laws and regulations of the Great Depression era, promoted the unionization of unskilled workers. As workers gained stronger bargaining position and more clout, they were able to influence the type of technology used by employers. Similarly, reasons unrelated to technology that led to the decline of unions, such as waning political support for unions and a roll back of laws favorable to labor, allowed employers to move away from unskilled labor-intensive production methods. This alternative hypothesis does not find as much support in historical evidence.

First, the alternative hypothesis does not line up well with the timing of events. The relative timing of the rise of mass production and unionization, examined in a variety of ways in Section 2.1, indicates that the rise of unions started well after the first emergence and initial diffusion of mass production in the 1920s. The alternative hypothesis, that mass production technologies were adopted purely as a consequence of an exogenous increase in the membership and power of unions, would have difficulty explaining this timing. Similarly, the steep decline of unions occurred mainly in the 1970s and the 1980s, after the emergence and early diffusion of technologies that favored skilled, as opposed to unskilled, labor. This timing was discussed in Section 2.2.

Second, the alternative hypothesis is also at odds with the theories of technology choice by firms. It cannot explain why businesses continued to adopt mass production technologies at an accelerating pace during the 1930s, and expanded their output and employment, even as unions continued to make sharp gains in membership. Suppose that there was no change in the relative productivity of unskilled workers, but increasingly more powerful unions pushed for higher employment or wages for such workers. Faced with the prospect of having to employ a large number of union workers at a higher wage with no productivity gain, businesses would have looked for ways to substitute away from unskilled-labor-intensive mass production methods. The evidence points against such an outcome: the diffusion of mass production, and the expansion of employment and output in mass production industries, continued rapidly after the surge in unionization in the early 1930s, as documented in Figures 2 and 3. It must have still paid off for firms to continue employing unionized labor and expand production, despite the potentially higher cost of doing so.

Third, the experience of unions in coal mining provides further evidence against the alternative hypothesis. Unions in coal mining had a very different experience compared with unions in mass production industries. They are the exception that proves the rule. The United Mine Workers (UMW) union, the main union for coal miners, enjoyed large gains in membership and power before and during World War I. However, while most of the industrial unions thrived during the 1920s and 1930s, the UMW experienced a sharp decline during this period. This decline is inconsistent with a hypothesis that the union-friendly environment of the 1930s was the main force in the rise of unions. Under this view, the UMW should also have benefitted from this conducive environment. It did not, because the technological developments in coal mining during this period were unfavorable to unskilled coal miners. As noted by Fox (1990) in his detailed history of the UMW, the main reason behind the decline in the UMW during the period was the diffusion of efficient coal-cutting machines. Before 1920, less than 41% of coal was cut by such machines. By 1930, nearly 81% was being cut by machines, and now new machines could also surface-mine and load the coal for transportation. As machines displaced unskilled labor, unemployment in the mines increased and wages were depressed. The decline of the UMW continued throughout the 1930s, unlike the industrial unions which grew substantially. The divergent patterns exhibited by UMW and the industrial unions during this period more readily fit an explanation based on the differential effects of technological change on unskilled labor in the two sectors. In the case of coal mining, technological change appears to have been biased against unskilled labor, whereas in mass production industries it was biased toward unskilled labor. Of course, in the second half of the 20th century many industrial unions shared the fate of the UMW in the 1920s and 1930s, as technology this time worked against unskilled workers.

Finally, the laws and regulations that shifted first in favor of unions during the 1930s, and against them later, can also be viewed as a consequence, rather than the cause, of the rise and fall of unions. Changes in labor laws and regulations undoubtedly contributed to the exceptionally rapid rise, and then the fall, of unionization. The shift in labor laws during the 1930s is chronicled in Ohanian (2009). Union wages were required to be paid on federal public works contracts by the Davis–Bacon Act in 1931. The Norris–Laguardia Act, which was enacted in 1932, limited the power of courts to issue injunctions against union strikes, picketing, or boycotts. It also outlawed "yellow dog" contracts. These contracts prohibited workers from joining a union; they could be fired if they did. The Wagner Act of 1935 provided for collective bargaining and placed very few restrictions on the rights of workers to strike. However, the dawning of the mass production era may have been a catalyst for enacting such laws. The growing demand for unskilled workers and their push to unionize under mass production may have resulted in the critical mass and political clout necessary for the lawmakers to act in their favor. Laws and regulations were also influential in the decline of the unions. Most importantly, some of the rights that unions had won during the 1930s were rolled by back by the Taft-Hartley Act in 1947. It outlawed closed shops, required an 80-day notice for strikes, and allowed states to pass "right-to-work" laws. The timing of the Act in 1947, which coincides with one of the high-points of union membership, is consistent with a growing reaction in the business community to the rapid diffusion of unions. However, the steep decline of unions did not really start until the 1970s, as seen in Figure 1.

3. The Setting

Imagine a world inhabited by a representative family with tastes given by

$$\sum_{t=1}^{\infty} \beta^{t-1} \ln \mathbf{c}_t, \text{ with } 0 < \beta < 1, \tag{1}$$

where $\mathbf{c}_t$ represents household consumption in period t. The family is made up of a continuum of members with a mass of one. Each household member supplies one unit of labor. A fraction

σ of these members are skilled, the rest unskilled.[12] A skilled worker earns the wage rate v_t. Unskilled members may work in the unionized part of the labor force or in the non-unionized one. A unionized worker earns the wage rate u_t, while a non-unionized one receives w_t. The fraction of unskilled household members that work during period t in the unionized part of the labor force is m_t, a variable that is determined in equilibrium. The household saves in the form of physical capital, $\mathbf{k}_t$. A unit of physical capital earns the rental r_t, and depreciates over time at rate δ. Finally, the household earns profits, $\boldsymbol{\pi}_t$, from the firms that it owns.[13]

There is a unit mass of firms in the economy. In period t a firm produces output, o_t, according to the production function

$$o_t = x_t z k_t^{\kappa}[\theta_t l_t^{\rho} + (1-\theta_t)(\xi_t s_t)^{\rho}]^{\alpha/\rho}, \text{ with } 0 < \alpha + \kappa < 1, \tag{2}$$

where k_t represents the amount of capital hired, l_t denotes the input of unskilled labor and s_t is the quantity of skilled labor. The variable x_t is a neutral shift factor for the technology that is common across firms. The variable x_t is assumed to grow at the constant rate g; $x_t = x_0 g^t$. A firm-specific shift factor is given by $z > 1$. This denotes a firm's type and is drawn at the beginning of time from a Pareto distribution

$$z \sim F(z) \equiv \frac{\zeta}{z^{\zeta+1}}, \text{ for } z > 1, \tag{3}$$

where F is the density function for a Pareto distribution.

Observe that skilled and unskilled labor are aggregated via a CES production function. The technology variables θ_t and ξ_t change over time and will capture the notion of skill-biased technological change.[14] There are diminishing returns to scale in production (since $\alpha + \kappa < 1$). There is a fixed cost ϕ_t associated with operating a firm. This fixed cost is assumed to grow at a constant rate, $\phi_t = \phi_1 (g^t)^{1/(1-\kappa)}$–the same rate at which output and wages grow along a balanced growth path in the model. The combination of diminishing

[12] The relative supply of skilled versus unskilled labor is assumed to be fixed over time. There is no doubt that supply shifts have occurred over the course of history, in particular due both to changes in the return from and the cost of an education (the latter due to changes in its public provision). The model abstracts from these supply effects.

[13] Variables in bold represent economy-wide aggregates. Capital and profits at the aggregate level need to be distinguished from their analogous firm-level quantities.

[14] It is uncommon to let θ_t vary over time. Rios-Rull and Santaeulalia-Llopis (2010) use a similar approach in their study of how labor's share of income fluctuates over the business cycle.

returns to scale in production and a fixed operating cost ensure that it is not desirable to organize all the firms in the economy.

Finally, there is a union in the economy. The union organizes unskilled labor in firms. An organized firm must use union workers for unskilled labor. The union believes in equality so all union members are paid the same wage, u_t. Unionization is a costly activity. Specifically, the period-t cost of organizing is given by the quadratic function

$$\chi_t \frac{m_t^2}{2}, \tag{4}$$

where m_t is the number of union members. The parameter χ_t grows in line with $\chi_t = \chi_1 (g^t)^{1/(1-\kappa)}$. The costs of organizing are recovered from the members in the form of dues, d_t. Skilled labor is not unionized. In the real world, this may be because skilled labor is too heterogenous in nature to be organized effectively to bargain for a common wage. The union is given the following set of preferences:

$$\sum_{t=1}^{\infty} \beta^{t-1} (u_t - d_t - w_t)^{\omega} m_t^{1-\omega}, \text{ with } 0 < \beta, \omega < 1. \tag{5}$$

These preferences presume that the union has two regards. It values the surplus that a union member earns over a non-unionized worker, $u_t - d_t - w_t$, as well as the number of unionized workers, m_t, that receive it. As will be seen, there is a trade-off between these two regards.

4. Decision Problems

The problems of the households and the firms in the economy are now described in detail, followed by the problem of the union.

4.1. Households

The problem facing the representative family is standard, with due alteration for the setting under study. Specifically, the household desires to maximize its lifetime utility subject to the budget constraint it faces each period:

$$\max_{\{\mathbf{c}_t, \mathbf{k}_{t+1}\}_{t=1}^{\infty}} \sum_{t=1}^{\infty} \beta^{t-1} \ln \mathbf{c}_t, \tag{6}$$

subject to

$$\mathbf{c}_t + \mathbf{k}_{t+1} = (1 - \sigma - m_t)w_t + m_t(u_t - d_t) + \sigma v_t + (r_t + 1 - \delta)\mathbf{k}_t + \boldsymbol{\pi}_t, \text{ for all } t.$$

In the above maximization problem the household takes the number of union members, m_t, as given. Since $u_t - d_t > w_t$, it would like as many unskilled household members as possible to be employed in union firms.

4.2. Firms

A firm in period t hires capital, k_t, and skilled and unskilled labor, s_t and l_t, to maximize profits. The firm's period-t choice problem is

$$\Pi_t^q(z; q_t, \cdot) \equiv \max_{k_t^q, l_t^q, s_t^q} \{x_t z (k_t^q)^\kappa [\theta_t (l_t^q)^\rho + (1 - \theta_t)(\xi_t s_t^q)^\rho]^{\alpha/\rho} - r_t k_t^q - v_t s_t^q - q_t l_t^q\} - \phi_t, \quad (7)$$

for $q = u, w$. With some abuse of notation, the variable q in superscript form will denote whether the firm is unionized ($q = u$) or not ($q = w$), while the variable q in regular form will represent the wage rate (again for $q = u, w$). Now, express the solution to the above problem for the amount of unskilled labor that a type-z firm will hire at the wage rate q_t by $l_t^q(z) = L_t^q(z; q_t, \cdot)$, for $q = u, w$–the "$\cdot$" represents the other arguments that enter the function L^q, which are suppressed to keep the subsequent presentation simple. Likewise, represent the amount of capital and skilled labor hired by $k_t^q(z) = K_t^q(z; q_t, \cdot)$ and $s_t^q(z) = S_t^q(z; q_t, \cdot)$. A firm's output is denoted by $o_t^q(z) = O_t^q(z; q_t, \cdot)$ and its profits are written as $\pi_t^q(z) = \Pi_t^q(z; q_t, \cdot)$.

Production is not a foregone conclusion due to the presence of the fixed operation cost, ϕ_t. A firm will only produce if it makes nonnegative profits. Thus, it must transpire that in equilibrium

$$\pi_t^q(z) = \Pi_t^q(z; q_t, \cdot) \geq 0, \text{ for } q = u, w. \quad (8)$$

Denote the period-t threshold value for z, at which it is just profitable for a firm to produce, by z_t^q. This threshold value solves the equation

$$\pi_t^q(z_t^q) = \Pi_t^q(z_t^q; q_t, \cdot) = 0, \text{ for } q = u, w. \quad (9)$$

It should be clear that $\Pi_t^q(z_t^q; q_t, \cdot) > 0$ for $z > z_t^q$ and $\Pi_t^q(z_t^q; q_t, \cdot) < 0$ for $z < z_t^q$.

From the two first-order conditions associated with hiring labor, it transpires that

$$\frac{s_t^q}{l_t^q} = [\xi_t^\rho \frac{(1-\theta_t)}{\theta_t} \times \frac{q_t}{v_t}]^{1/(1-\rho)}, \text{ for } q = u, w. \tag{10}$$

The ratio of skilled to unskilled labor, s_t^q/l_t^q, depends on the term that captures the price of unskilled labor relative to skilled labor, $(q_t/v_t)^{1/(1-\rho)}$. It also depends on the skill-biased technology term, $\epsilon_t \equiv [\xi_t^\rho(1-\theta_t)/\theta_t]^{1/(1-\rho)}$, which captures the notion of skill-biased technological change. The benefit of unionizing unskilled workers is large when ϵ_t is small, because unskilled labor is favored, relatively speaking. The skill-biased term, ϵ_t, will be low when the weight on unskilled labor in production, θ_t, is high. How the productivity of skilled labor, ξ_t, affects the skill-bias term, ϵ_t, depends on the sign of ρ. When skilled and unskilled labor are more substitutable than in the case of Cobb-Douglas production function, or when $\rho > 0$, a drop in the productivity of skilled labor, ξ_t, will reduce the skill-bias term, ϵ_t. When they are less substitutable, $\rho < 0$, things go the other way. Last, note that the production function (2) can be rewritten as $o_t = x_t z k_t^\kappa \theta_t^{\alpha/\rho}[l_t^\rho + \epsilon_t^{1-\rho} {s_t}^\rho]^{\alpha/\rho}$. Thus, the profitability of firms, which affects the extent of unionization, is a function of ϵ_t and θ_t. This observation is useful because it implies that θ_t and ϵ_t, and hence ξ_t, are all identified in the quantitative analysis by using data on the income distribution and the extent of unionization.

4.3. The Union

Recall that the union has two regards. First, it values the surplus over the competitive wage that union members earn. Second, it also puts worth on the number of workers that will earn the union wage. The union organizes the firms with the highest levels of productivity first, because they can better afford to pay the union premium and provide larger union employment. There is, however, a limit to the wage that the union can set. Specifically, a unionized firm must earn nonnegative profits. So, if any unionized firm earns zero profits, then all firms with a higher level of productivity will be unionized and those with a lower level will not. Because more productive firms are also larger in the model, the union organizes larger firms.[15]

Now, turn to the optimization problem faced by a union. Assume that the profits of the last firm unionized are squeezed to zero. The number of unionized workers in period t, m_t, will be given by

$$m_t = \int_{z_t^u}^{\infty} L_t^u(z; u_t, \cdot)F(z)dz. \tag{11}$$

The dues paid by a union member, d_t, are

$$d_t = \frac{\chi m_t^2}{2m_t} = \frac{\chi_t[\int_{z_t^u}^{\infty} L_t^u(z; u_t, \cdot)F(z)dz]}{2}. \tag{12}$$

The union's decision problem appears as

$$\max_{\{u_t, z_t^u\}_{t=1}^{\infty}} \sum_{t=1}^{\infty} \{u_t - \frac{\chi_t[\int_{z_t^u}^{\infty} L_t^u(z; u_t, \cdot)F(z)dz]}{2} - w_t\}^{\omega}[\int_{z_t^u}^{\infty} L_t^u(z; u_t, \cdot)F(z)dz]^{1-\omega}, \tag{13}$$

subject to the zero-profit constraint (9) holding (when $q = u$) for the marginal union firm, z_t^u. When solving its problem, the union takes the wages for non-unionized unskilled and skilled labor, w_t and v_t, as given.

To understand why the union will pick the wage rate so that the threshold firm earns zero profits, suppose that the marginal firm earned positive profits. The cost of raising the union wage, u_t, incrementally is the loss of membership that will occur from all of the inframarginal firms. It turns out, though, that this loss can be made up for by increasing the number of unionized firms or lowering z_t^u. The process of simultaneously raising u_t and lowering z_t^u cannot go on forever. At some point, the firms with the lowest z_t^u will no longer be able to earn profits due to the presence of the fixed cost ϕ_t. Then, the process must stop.

Lemma 1 *(Zero profits for the marginal firm) The union picks the wage rate, u_t, so that the zero-profit constraint (9) is binding (when $q = u$) for the last firm organized.*

[15] Consistent with this prediction, Dinlersoz, Greenwood, and Hyatt (2016) find a higher likelihood of union activity among more productive and larger businesses. Specifically, they match data on union elections in business establishments with the data on establishment characteristics for the entire set of establishments in the United States during the period 1977-2007. The analysis indicates that: (i) larger, more productive and younger establishments are more likely to experience a union certification election; (ii) unions have a higher probability of being certified in larger and younger establishments; and (iii) union contracts are more prevalent in larger, more productive and older establishments. See also Fang and Heywood (2006) for the connection between unionization and plant size in the case of Canada.

Proof. See the Online Appendix. ■

The union's two regards must be traded off in the maximization problem (13). By applying the envelope theorem to a unionized firm's optimization problem (7), for $q = u$, it can be easily calculated from equation (9) that

$$\frac{du_t}{dz_t^u} = \frac{O^u(z_t^u; u_t, \cdot)}{z_t^u L_t^u(z_t^u; u_t, \cdot)} > 0. \tag{14}$$

This implies that lowering the threshold hold, z_t^u, or equivalently unionizing more firms, can only be accomplished by reducing the union wage, u_t. Additionally, it can be seen from equation (12) that a rise in membership, $m_t = \int_{z_t^u}^{\infty} L_t^u(z; u_t, \cdot)F(z)dz$, comes at the expense of higher dues, d_t, because of the increasing costs involved with unionization.

5. Equilibrium

In equilibrium the markets for capital, labor and goods must clear. Equilibrium in the capital market requires that

$$\int_{z_t^w}^{z_t^u} k_t^w(z)F(z)dz + \int_{z_t^u}^{\infty} k_t^u(z)F(z)dz = \mathbf{k}_t. \tag{15}$$

The market-clearing condition for skilled labor is

$$\int_{z_t^w}^{z_t^u} s_t^w(z)F(z)dz + \int_{z_t^u}^{\infty} s_t^u(z)F(z)dz = \sigma, \tag{16}$$

while that for unskilled labor reads

$$\int_{z_t^w}^{z_t^u} l_t^w(z)F(z)dz + \int_{z_t^u}^{\infty} l_t^u(z)F(z)dz = 1 - \sigma. \tag{17}$$

Last, equilibrium in the goods market implies

$$\mathbf{c}_t + \mathbf{k}_{t+1} + m_t d_t + \phi_t = \int_{z_t^w}^{z_t^u} o_t^w(z)F(z)dz + \int_{z_t^u}^{\infty} o_t^u(z)F(z)dz + (1-\delta)\mathbf{k}_t. \tag{18}$$

Note that the aggregate amount of union dues, $m_t d_t$, appears in the resource constraint. These exactly cover the resource cost of organizing–see (12).

A definition of the equilibrium under study will now be presented to take stock of the situation so far.

Definition 2 *(Definition of a competitive equilibrium) A competitive equilibrium is a time path for consumption and savings,* $\{\mathbf{c}_t, \mathbf{k}_{t+1}\}_{t=1}^{\infty}$*, a set of labor and capital allocations for union* ($q = u$) *and non-union* ($q = w$) *firms* $\{l_t^q(z), s_t^q(z), k_t^q(z)\}_{t=1}^{\infty}$*, a set of factor prices* $\{u_t, w_t, v_t, r_t\}_{t=1}^{\infty}$*, a sequence for union dues,* $\{d_t\}_{t=1}^{\infty}$*, and a sequence determining the threshold points for union and non-union firms,* $\{z_t^u, z_t^w\}_{t=1}^{\infty}$*, such that for a given time profile for technology* $\{\theta_t, \xi_t, x_t, \phi_t, \chi_t\}_{t=1}^{\infty}$ *the following five conditions hold:*

First, the time path for consumption and savings, $\{\mathbf{c}_t, \mathbf{k}_{t+1}\}_{t=1}^{\infty}$*, solves the representative household's problem (6) given the time path for factor prices,* $\{u_t, w_t, v_t, r_t\}_{t=1}^{\infty}$*, profits,* $\boldsymbol{\pi}_t \equiv \int_{z_t^w}^{z_t^u} \pi_t^w(z)F(z)dz + \int_{z_t^u}^{\infty} \pi_t^u(z)F(z)dz$*, and the size of the union sector,* $m_t = \int_{z_t^u}^{\infty} l_t^u(z)F(z)dz$*.*

Second, the time paths for firms' input utilizations, $\{l_t^q(z), s_t^q(z), k_t^q(z)\}_{t=1}^{\infty}$*, solve their profit maximization problems, as specified by (7), given the time paths for factor prices,* $\{q_t, v_t, r_t\}_{t=1}^{\infty}$ *(for* $q = u, w$*) and technology* $\{\theta_t, \xi_t, x_t, \phi_t, \chi_t\}_{t=1}^{\infty}$*.*

Third, the sequences for the union wage, $\{u_t\}_{t=1}^{\infty}$ *and the threshold,* $\{z_t^u\}_{t=1}^{\infty}$*, solve the union's problem (13), given the time paths for factor prices,* $\{q_t, v_t, r_t\}_{t=1}^{\infty}$ *(for* $q = u, w$*), technology,* $\{\theta_t, \xi_t, x_t, \phi_t, \chi_t\}_{t=1}^{\infty}$*, and the solution to the unionized firm's problem,* $l_t^u(z) = L_t^u(z; u_t, \cdot)$ *and* $\pi_t^u(z) = \Pi_t^u(z; u_t, \cdot)$*, as implied by (7). The sequence for union dues,* $\{d_t\}_{t=1}^{\infty}$*, is determined in line with (12).*

Fourth, the sequence for the threshold, $\{z_t^w\}_{t=1}^{\infty}$*, solves (9) when* $q = w$*, given* $\pi_t^w(z) = \Pi_t^w(z; w_t, \cdot)$ *from (7), and the series* $\{w_t, v_t, r_t\}_{t=1}^{\infty}$ *and* $\{\theta_t, \xi_t, x_t, \phi_t, \chi_t\}_{t=1}^{\infty}$*.*

Fifth, the markets for capital, labor and goods all clear, so that equations (15) to (18) hold.

6. Simulation Analysis

The model described above is now simulated to assess its ability to replicate the time paths for union membership and income inequality in the United States over the course of the 20th century.

6.1. Calibration

Before the model can be simulated, values must be assigned for its parameters. Table 2 lists the parameter values. The period length is five years. Some parameters can be chosen on the basis of a priori information. Accordingly, the discount factor is set so that $\beta = 1/(1.04)^5 \simeq 0.82$, which implies an annual steady-state interest rate of 4%, a standard value. The annual depreciation rate for capital is taken to be 0.08, another standard value. Likewise, labor's share of income is set at 60%, implying $\alpha = 0.60$, a typical value if one assumes that part of the capital stock includes intangibles–see Corrado, Hulten and Sichel (2009). The production function for firms exhibit diminishing returns to scale. Guner, Ventura and Xi (2008) report that the share of profits in output is 20%. Capital's share of income, κ, is therefore set at 0.20. Katz and Murphy (1992) estimate that the elasticity of substitution between skilled and unskilled labor is 1.4, corresponding to a value of 0.29 for ρ.

The rest of the model's parameters are selected using a calibration procedure. The procedure requires that a steady state for the model hits 5 data targets for the year 1955, roughly the peak year of the unionization movement–see Figure 1. This involves computing the model's steady state in conjunction with the 5 data targets, while taking the 5 parameters θ_{1955}, ϕ, ω, χ and ζ as additional variables. The technology variable ξ is normalized so that $\xi_{1955} = 1$. While the nonlinear system of equations used to calibrate the model is simultaneous in nature, certain parameters play a key role in matching each of the data targets. The five data targets and their importance for identifying the five parameters are discussed now.

The first target is the union membership rate of 37% in 1955. Therefore, the steady state is computed subject to the restriction $m_{1955} = 0.37$. The weight on the extent of membership in the union's objective function, $1 - \omega$, plays a key role in attaining this target (strictly speaking, the other 4 parameters play some role in achieving this too). Thus, one can think about the level of union membership in 1955 as identifying the parameter ω.

Let the top 10% of the population represent skilled labor. Thus, $\sigma = 0.10$. The use of 10% is a compromise. On the one hand, using a smaller value, such as 5% or 1%, risks including individuals whose earnings and income may depend on sources other than skilled-labor's earnings. Using a larger value, on the other hand, leads to the inclusion of many unskilled individuals who are in the right tail of the income distribution for unskilled labor.[16] The share of the top 10% of the work force in earnings was 0.32 in 1955. Therefore, the steady state must satisfy the restriction $\sigma v_{1955}[m_{1955}u_{1955} + (1 - m_{1955} - \sigma)w_{1955} + \sigma v_{1955}] = 0.32$. Not surprisingly, the constant term on unskilled labor in the production function, θ_{1955}, is very important for hitting this second target (although again the other 4 parameters impinge on this objective as well). So, heuristically speaking, θ_{1955} is identified from the share of the top 10% in the earnings distribution in 1955.

Union dues are assumed to amount to 1% of a union member's wages. MacDonald and Robinson (1992, p. 47) state that this is a reasonable value. Indeed, this value has stayed relatively constant over time, and it is exactly what the UAW currently charges salaried workers. Thus, the third target can be expressed as $d_{1955}/u_{1955} = 0.01$. The term χ in the union's cost function is instrumental in meeting this target.

The fourth target describes the firm size distribution. In the model, there is no distinction between establishments and firms. However, in the U.S. unionization predominantly occurs at the establishment level, so the distribution of employment at the establishment level is the focus. This distribution is highly skewed. Based on the *Longitudinal Business Database* of the U.S. Census Bureau, the coefficient of variation (CV) of employment across all U.S. establishments with at least one employee had an average value of approximately 7 over the period 1976-2011, varying in a narrow band of 6 to 8.[17] Comprehensive data for earlier years is difficult to find for this statistic. The earliest available data is for all manufacturing establishments in the 1963 Census of Manufactures, which reveals a CV of approximately 6 for employment. This value falls in the range of values calculated for 1976-2011. The average value of 7 guides the choice of the Pareto distribution parameter, ζ. The fourth target is

[16] For instance, according to Hirschl and Rank (2015), 61% of U.S. households break into the 20th percentile of the income distribution for at least two consecutive years. The share of the top 20% in the income distribution still has a $\cup$ shape, but it is a little less pronounced. Figure 4 shows that a wide variety of measures for income inequality have similar $\cup$-shaped patterns. Thus, the $\cup$-shaped pattern used in the analysis appears to be quite robust for different choices of σ.

[17] See the Online Appendix for the calculation of the descriptive statistics for the establishment-size distribution.

represented by $\text{CV}(l+s)=7$.

Last, in classic work, H. Greg Lewis (1963) reported that the union wage premium was 15%. Here, Card, Lemieux, and Riddell's (2003, Table 4) more recent estimate of 20% is used, giving the fifth target $u_{1955/}w_{1955} = 1.20$ (Using the Lewis's number instead does not make a difference for the results). Remember that the last firm unionized earns zero profits when it pays the union wage, u_t, and incurs the fixed cost, ϕ_t. The above condition can be used to back out the fixed cost in 1955, ϕ.

6.2. Results

The model's ability to account for the $\cap$-shaped pattern of union membership along with the $\cup$-shaped profile for income inequality over the 20th century is assessed now. To do this, the model's transitional dynamics are computed for the years 1925 to 2000. Truncating a few years from the available series for union membership that starts in 1910 avoids the temporary jump and drop in unionization associated with World War I, about which the analysis has nothing to say. Computing the transitional dynamics requires inputting in a time series process for technology, $\{\theta_t, \xi_t\}_{t=1925}^{2000}$.[18] Before undertaking the transitional dynamics, three preparatory steps are taken.

Fitting long-run trends for the income distribution and union membership. In the first step, quartic polynomials are fit to the income distribution and union membership time series to capture the long-run trends in both series. These polynomials go exactly through the data points for the five years 1925, 1945, 1955, 1980, and 2000. This step is taken because the goal of the analysis is to examine the long-run path of technological change, not the year-to-year fluctuations in these two series. In particular, the first half of the 20th century was quite tumultuous. It witnessed the Second Industrial Revolution, the Great Depression, and World Wars I and II. It would not be desirable (and even be believable) to have the model explain all the fluctuations associated with these events. A similar caveat applies to the

[18] A growth-transformed version of the model is simulated. This essentially amounts to assuming that x is constant or that $g = 1$. The model is described by a nonlinear difference equation system. Solving this system amounts to computing the saddle path solution for a two-point boundary value problem. The first boundary condition for the economy is the initial capital stock, while the second one is capital stock associated with the terminal steady state.

second half of the century. As can be seen from Figure 6, these quartics capture the long-run patterns quite well. The quartic fits for the time series for income distribution and union membership are used solely as a reference to compare the model's output with.

Backing out the technology parameters, θ_t and ξ_t, *for the five years 1925, 1945, 1955, 1980, and 2000.* Recall that values for θ_t and ξ_t for the year 1955 have already been determined. For each of the four remaining years, 1925, 1945, 1980, and 2000, steady states for the model are computed, taking union membership and income inequality as targets. Values of θ_t and ξ_t that hit these targets exactly are then backed out in the second step. Since θ_t and ξ_t enter into the skill-biased term, $\epsilon_t \equiv [\xi_t^{\rho}(1-\theta_t)/\theta_t]^{1/(1-\rho)}$, they both influence the income distribution. They also enter profits through the production function (2), and hence, affect the incentives for unionization.

Filling in the rest of the process for $\{\theta_t, \xi_t\}_{t=1925}^{2000}$. For the third step, assume that the time paths for θ_t and ξ_t are described by two quartics, each of which is fit to the five points previously obtained, $(\theta_{1925}, \theta_{1945}, \theta_{1955}, \theta_{1980}, \theta_{2000})$ and $(\xi_{1925}, \xi_{1945}, \xi_{1955}, \xi_{1980}, \xi_{2000})$, respectively. After the year 2000 all technological change is shut off.

The capital stocks associated with the 1925 and 2000 steady states are taken as the initial and terminal capital stocks. Union membership and income inequality are then calculated every five years for the period 1925 to 2000 along the path for the transitional dynamics.

The framework does an excellent job in accounting for the rise and fall in the quartic fit for union membership, as the right panel of Figure 6 illustrates. The circles on the diagram show the five points that the quartic for union membership is fit to. The model ever so slightly overshoots union membership at the quartic point for 1980. It also mimics the fall and rise in the quartic fit for income inequality fairly well too. This is shown in the left panel of the figure, where the circles show the five points for income inequality that the quartic was fit to. Observe that the model underpredicts the high level of inequality in 2000. This is because the model's transitional dynamics have not quite converged yet to the values for 2000 steady state. In the analysis, skill-biased technological change is the sole driver of both the time series for unionization and income inequality; i.e., the $\cup$-shaped pattern in income inequality is not caused by the $\cap$-shaped time series for unionization. By this account, very little of postwar rise in inequality can be accounted for by the decline in unionization.[19] Goldin and Katz (2008), Greenwood and Yorukoglu (1997), and Krusell et al. (2000) all stress technological change as a force underlying shifts in the income distribution.

6.3. Skill-Biased Technological Change

One could ask whether the constructed time series for technology that is inputted into the simulation, $\{\theta_t, \xi_t\}_{t=1925}^{2000}$, is reasonable or not. The resulting time profile for skill-biased technological change, shown in the right panel of Figure 7, is $\cup$-shaped. The evolution of ϵ_t favors unskilled labor during the first part the century, in agreement with the diffusion of mass production in Figures 2 and 3. After 1955 the skill-biased term starts to benefit skilled labor. Consistent with this path, Figure 3 indicates a rise in "computers," "automation" and "robots," and a fall in "mass production." Moreover, the relative price of equipment and software fell rapidly over this period and the share of information processing equipment in private investment took off, as Figure 5 illustrates. So, the pattern for skill-biased technological change in Figure 7 conforms qualitatively with what is observed in the available measures.

The values of θ_t and ξ_t associated with the skill-biased technological change series are also in the left panel of Figure 7. The squares on the graph show the five points that the quartics for θ_t and ξ_t were fit to. These five points for each of the two technology variables are backed out by forcing the model's steady state to hit the five targets (each) for the income distribution and membership series shown by the circles in the left panel of Figure 6.[20] The weight on unskilled labor, θ_t, is $\cap$-shaped. One might have expected the term augmenting skill, ξ_t, to be $\cup$-shaped. The very different patterns for θ_t and ξ_t speak to the fact that they

[19] A drop in unionization has a minor impact on income inequality in the model. This is due to the facts that (i) the union wage premium is of moderate size, and (ii) it applies to a relatively small part of the aggregate wage bill. In fact, if one assumes that all unskilled workers get the non-union wage then the plot obtained for the income distribution looks virtually identical to that displayed in Figure 6.

[20] It is reasonable to ask how union dues, the coefficient of variation of employment across firms, and the union wage premium behave in the steady states for these five dates. The results for 1925, 1945, 1955, 1980 and 2000 are: 3%, 1%, 1%, 0.8%, and 0.8% for union dues; 5.9, 7.1, 7.0, 6.3, and 5.8 for the coefficient of variation; and 1.23, 1.20, 1.20, 1.20, and 1.22 for the union wage premium. A small fraction of firms (<3%) in the 2000 steady state earn negative profits. This is because the constraint for nonnegative profits for non-union firms is ignored in the simulation for computational convenience.

are separately identified. A decline (increase) in the skill-bias term, ϵ_t, can be obtained in two ways: lowering (raising) θ_t or raising (lowering) ξ_t. Loosely speaking, the movement in the skill-bias term, ϵ_t, which could be obtained either through θ_t and ξ_t, is identified from the observed shift in income inequality. The terms θ_t and ξ_t affect the incentive to unionize differently. Recall that output can be written as $o_t = x_t z k_t^{\kappa} \theta_t^{\alpha/\rho} [l_t^{\rho} + \epsilon_t^{1-\rho} s_t{}^{\rho}]^{\alpha/\rho}$. Hence, for a given ϵ_t, an increase in θ_t raises the profitability of production and spurs unionization. Therefore, to obtain a large jump in unionization in the first half of the century, the model likes a big step up in θ_t or a large drop in $(1-\theta_t)/\theta_t$. Matching the observed skill premium then requires a compensating rise in ξ_t. Note that the upward movement in ξ_t does *not* mean that skilled labor became more productive in the first part of the century and less productive in the second. The relative productivity of skilled labor is governed by the *combined* effects of θ_t and ξ_t in the skill-bias term ϵ_t, which exhibits the expected $\cup$-shaped pattern.

The magnitude of skill-biased technological change in Figure 7 is reasonable. The skill-biased term drops by a factor of 1.5 from 1955 to 2000. Over the 1925 to 2000 period real per-capita income grew by 2.25% a year. This implies that real per-capita GDP rose by a factor of 5.3. To achieve this in the model, the parameter governing neutral technological change, x_t, must rise by a factor of $g^{2000-1925} = 5.3^{1-\kappa} = 3.8$, which implies $g = 1.018$, or that x_t grows at 1.8% a year. Therefore, the required amount of skill-biased technological change is smaller than that of neutral technological change, as Figure 6 illustrates. A crude measure of the realism of the magnitude of skill-biased technological change is to take (10) to the U.S. data. Over the time period 1958 to 2000 the ratio of skilled-to-unskilled labor, as proxied by the ratio of non-production-to-production workers in U.S. manufacturing, rose by a factor of 1.22, while the skill premium, measured by the ratio of the average earnings of these workers, increased by 1.15.[21] When $\rho = 0.29$, as used in the calibration (Table 2), these figures imply that the skill-biased term moved up by a factor of 1.4, which is close to the number estimated here, 1.5, based on the growth of the skill-biased term in Figure 7 between 1958 and 2000.

Last, direct measures of skilled-biased technological change at the plant or firm level are

[21] The skill ratio and premium are calculated using the NBER-CES Manufacturing Database available at http://www.nber.org/nberces/ (Last accessed: September, 2015).

hard to come by. Research in progress aims to overcome this shortcoming by using the U.S. Census Bureau's 1991 Survey of Manufacturing Technology. The research uses plant-level microdata data for selected manufacturing industries on technology adoption and use, as well as on the presence of a union contract for production workers. The analysis indicates that union contracts for production workers are less prevalent in plants with a higher intensity of technology adoption and use, after controlling for a range of other plant and industry characteristics. This result is especially strong in the case of fabrication and machining technologies. These technologies include robots, NC/CNC machines, and flexible manufacturing systems, which are generally highly substitutable with production labor. While this relationship should not be interpreted as a causal one, it supports the hypothesis that unions and computerized technologies do not go together. Therefore, the diffusion of advanced technologies in the economy may have adversely affected unionization, consistent with the model presented here.

6.4. Welfare Cost of Unions

The welfare cost of unions has been a question of long-standing interest in economics. Rees (1963) studied this question some time ago. He found that the welfare loss from unions in 1957 amounted to 0.14% of GDP. The model developed here can also be used to address this question. Suppose that the model economy is resting in its 1955 steady state, the peak of the union power. Now, eliminate unions. The model would then imply a welfare increase of 0.66% of GDP.[22] While this number is 4.7 times as big as Rees's, it is paltry. The welfare cost of unions is restricted here by the assumption that firms are competitive. Whether or not this is a good approximation for the U.S. economy over the time period studied is an open question. Perfect competition limits the wages that unions can obtain. Unions are more likely to have a large impact on economic activity when they are negotiating with producers that have monopoly power. This was the case in U.S. iron ore industry prior to the 1970s. After this time, producers faced intense competition from foreign exporters. Schmitz

[22] This welfare cost can also be computed using a more traditional Rees (1963) style diagram. See Figure A3 and the related explanation in the Online Appendix.

(2005) documents how this increased competition led to a large rise in labor productivity– see Greenwood and Weiss (2016) for a formal model of Schmitz's hypothesis. His analysis might also apply at points in time to the aircraft, airline, and auto industries, for example. Similarly, Cole and Ohanian (2004) study the impact that unions had on the economy during the Great Depression. They stress the cartelization of industries allowed by Roosevelt under the New Deal, which were then abandoned prior to World War II. Taschereau-Dumouchel (2015) argues that just the threat of unionization may be enough to generate large welfare costs. Finally, Alder, Lagakos and Ohanian (2015) suggest that union power played an important role in the decline of the rust belt.

7. Conclusion

The shift from an artisanal to mass production during the beginning of the 20th century was associated with an increase in the relative productivity of unskilled labor that led to an increase in unionization and a decrease in income inequality. The decline of mass production and the rise of the Information Age reversed this trend, leading to the $\cap$-shaped pattern of unionization and the $\cup$-shaped one for income inequality. To study the connection between technology, unionization and income inequality, a general equilibrium model of unionization is developed here. Heterogeneous firms hire capital, and skilled and unskilled labor. A union can organize unskilled labor at a cost. It cares about the wage rate that its members earn, and how many workers receive this wage. The union sets its membership and the wage so that it squeezes all of the rents from the last firm organized. The nature of technology influences the value of unskilled labor. When the productivity of unskilled labor is relatively high, it pays for the union to organize a lot of firms and demand generous wages.

The analysis proceeds on two fronts. First, historical evidence is presented to describe the evolution of unionization and the shifts in the mix of skilled and unskilled labor in the wake of some fundamental changes in the U.S. economy during the 20th century. It is argued that these changes were brought about by the introduction of mass production techniques in the first half of the century and by computerization in the second half. Second, the constructed model is calibrated and simulated to gauge whether it is capable of explaining the above stylized facts. It is. To obtain the patterns in data, the amount of skill bias must follow a $\cap$-shaped pattern. The required change in skill bias is plausible. It also mirrors the qualitative pattern expected from economic history.

References

Acemoglu, D., Aghion, P., Violante, G., 2001. Deunionization, Technical Change and Inequality. Carnegie-Rochester Conference Series on Public Policy, 55: 229-264.

Acikgoz, O., Kaymak, B., 2014. The Rising Skill Premium and Deunionization. Journal of Monetary Economics 63, 37-50.

Alder, S., Lagakos, D., Ohanian, L., 2015. Competitive Pressure and the Decline of the Rust Belt: A Macroeconomic Analysis. Manuscript, University of California at San Diego.

Alvaredo, F., Atkinson, T., Piketty, T., Saez, E., 2011. The World Top Incomes Database. Available at http://topincomes.g-mond.parisschoolofeconomics.eu/.

Card, D., Lemieux, T., Riddell, C., 2003. Unionization and Wage Inequality: A Comparative Study of the U.S., the U.K., and Canada. NBER Working Paper 9473.

Cole, H., Ohanian, L., 2004. New Deal Policies and the Persistence of the Great Depression: A General Equilibrium Analysis. Journal of Political Economy 112, 779-816.

Corrado C., Hulten, C.and Sichel, D., 2009. Intangible Capital and U.S. Economic Growth. Review of Income and Wealth 55, 661-685.

Cummins, J., Violante, G., 2002. Investment-Specific Technical Change in the United States (1947-2000): Measurement and Macroeconomic Consequences. Review of Economic Dynamics 5, 243-284.

Dinlersoz, E., Greenwood, J., Hyatt, H., 2016. What Businesses Attract Unions? Unionization over the Life-Cycle of U.S. Establishments. Industrial and Labor Relations Review, forthcoming.

Feng, T., Heywood, J. S., 2016. Unionization and Plant Closure in Canada. Canadian Journal of Economics 39, 1173-1194.

Finke, R., McClure, J. M., 2015. Reviewing Millions of Books: Charting Cultural and Religious Trends with Google's Ngram Viewer. Association of Religion Data Archive's Working Paper Series. (http://thearda.com/rrh/papers/book.asp)

Ford, H., 1926. Mass Production. In: Encyclopedia Britannica, 13th Edition, Volume 15, pp. 38-41.

Fox, M. B., 1990. United We Stand: The United Mine Workers of America, 1890-1990. Washington, DC: United Mine Workers of America.

French, M., 1992. Structural Change and Competition in the United States Tire Industry, 1920-1937. The Business History Review, 60: 28-54.

Goldin, C., Katz, L., 2001. Decreasing (and then Increasing) Inequality in America: A Tale of Two Half Centuries. In: Welch, F. The Causes and Consequences of Increasing Inequality. University of Chicago Press, pp. 37-82.

Goldin, C., Katz. L., 2008. The Race between Education and Technology. Harvard University Press, Massachusetts.

Greenwood, J., Hercowitz, Z., Krusell, P., 1997. Long-Run Implications of Investment-Specific Technological Change. American Economic Review 87, 342-362.

Greenwood, J. and Weiss, D. 2016. Mining Surplus: Modeling James A. Schmitz's Link Between Competition and Productivity. Unpublished Paper, Tel Aviv University.

Greenwood, J., Yorukoglu, M., 1997. 1974. Carnegie-Rochester Conference Series on Public Policy 46, 49-95.

Hampton, B. History of the American Film Industry: From its beginnings to 1931. Dover, New York.

Hirschl, T. A., Rank, M. R., 2015. The Life Course Dynamics of Affluence. PLoS ONE 10(1): e0116370. doi:10.1371/journal.pone.0116370.

Juhn, C., Murphy, K., Pierce, B., 1993. Wage Inequality and the Rise in Returns to Skill. Journal of Political Economy 101, 410-442.

Kahan, P., 2014. The Homestead Strike: Labor, Violence, and American Industry. Routledge, New York.

Katz, L., Murphy, K., 1992. Changes in Relative Wages, 1963–1987: Supply and Demand Factors. Quarterly Journal of Economics 107, 35–78.

Krusell, P., Ohanian, L., Ríos-Rull, J., Violante, G., 2000. Capital-skill Complementarity and Inequality: A Macroeconomic Analysis. Econometrica 68, 1029-1053.

MacDonald, G., Robinson, C., 1992. Unionism in a Competitive Industry. Journal of Labor Economics 10, 33-54.

McElvaine, R., 2009. The Great Depression. New York: Three Rivers Press.

Meyer, S., 1981. The Five Dollar Day: Labor Management and Social Control in the Ford Motor Company 1908-1921. State University of New York Press, Albany.

Meyer, S., 1988. Technology and the Workplace: Skilled and Production Workers at Allis-Chalmers, 1900-1941. Technology and Culture 29, 839-864.

Michel, J., Shen, Y., Aiden, A., Veres, A., Gray, M., The Google Books Team, Pickett, J., Hoiberg, D., Clancy, D., Norvig, P., Orwant, J., Pinker, S., Nowak, M., Aiden, E., 2011. Quantitative Analysis of Culture Using Millions of Digitized Books. Science 14, 176-182.

Nelson, D., 1987. Mass Production and the U.S. Tire Industry. Journal of Economic History 47, 329-339.

Nelson, D., 1988. American Rubber Workers and Organized Labor, 1900-1941. Princeton University Press, New Jersey.

Ohanian, L., 2009. What - or Who - Started the Great Depression? Journal of Economic Theory 144, 2310-2335.

Pease, W., 1952. An Automatic Machine Tool. Scientific American, September, 101-115.

Piore, M., 1986. The Decline of Mass Production and the Challenge to Union Survival. Industrial Relations Journal 17, 207-213.

Rees, A., 1963. The Effects of Unions on Resource Allocation. Journal of Law and Economics 6, 69-78.

Rios-Rull, J., Santaeulalia-Llopis, R., 2010. Redistributive Shocks and Productivity Shocks. Journal of Monetary Economics 57, 931-948.

Schmitz, Jr, J., 2005. What Determines Productivity? Lessons From the Dramatic Recovery of the U.S. and Canadian Iron Ore Industries Following Their Early 1980s Crisis. Journal of Political Economy 113, 582-625.

Solow, R., 1959. Investment and Technical Progress. In: Arrow, K., Karlin, S., Suppes, P. (Eds.), Mathematical Methods in the Social Sciences. Stanford University Press, Stanford, CA, pp. 89–104.

Storper, M., 1989. The Transition to Flexible Specialisation in the US Film Industry: External Economies, the Division of Labor, and the Crossing of Industrial Divides. Cambridge Journal of Economics 13, 273-305.

Taschereau-Dumouchel, M., 2015. The Union Threat. Manuscript. The Wharton School, University of Pennsylvania.

Williams, K., Haslam, C., Williams, J., 1992. Ford versus “Fordism”: The Beginning of Mass Production? Work, Employment, and Society 6: 517-555.

Wolmack, J., Jones, D., Roos, D., 1990. The Machine that Changed the World. Free Press, New York.

Zheng, R., Greenfield, P. M., 2015. Cultural Evolution over the Last 40 Years in China: Using the Google Ngram Viewer to Study Implications of Social and Political Change for Cultural Values. International Journal of Psychology 50, 47-55.

Table 1. Composition of workers in Detroit Metal Ind. and Ford Motor Co.

Occupation	**No.**	**Percent**	**Mean Weekly Income**
Detroit Metal Industries, 1891			
Foreman	9	2	$19.67
Mechanics	153	39	12.58
Specialists	117	30	8.18
Unskilled Labor	113	29	6.60
Total	392	100	9.55
Ford Motor Company, 1913			
Mechanics and Subforeman	329	2	–
Skilled Operators	3,431	26	–
Operators	6,749	51	–
Unskilled Workers	2,795	21	–
Total	13,304	100	–

Source: Meyer (1981, p. 46 and 50)

Table 2. Parameter Values

Selected using a priori information		
Parameter	**Definition**	**Source**
Tastes		
$\beta = (1.04)^{-5}$	discount factor	standard
Technology		
$\alpha = 0.60$	labor's share	Corrado et al. (2009)
$\delta = 1 - (1 - 0.08)^5$	depreciation rate	standard
$\kappa = 0.20$	exponent on capital	Guner et al. (2008)
$\rho = 0.29$	elasticity of substitution	Katz and Murphy (1992)
$\xi_{1955} = 1.0$	shift factor on skilled labor	normalization
Selected using the calibration procedure		
Parameter	**Definition**	**Basis of identification**
Technology		
$\theta_{1955} = 0.53$	weight on unskilled labor	skill premium
$\zeta = 2.0057$	Pareto distribution	establishment-size distribution
$\phi = 0.07$	fixed cost	union wage premium
Unionization		
$\omega = 0.36$	ideals–wage	union membership
$\chi = 0.01$	organization costs, constant	union dues

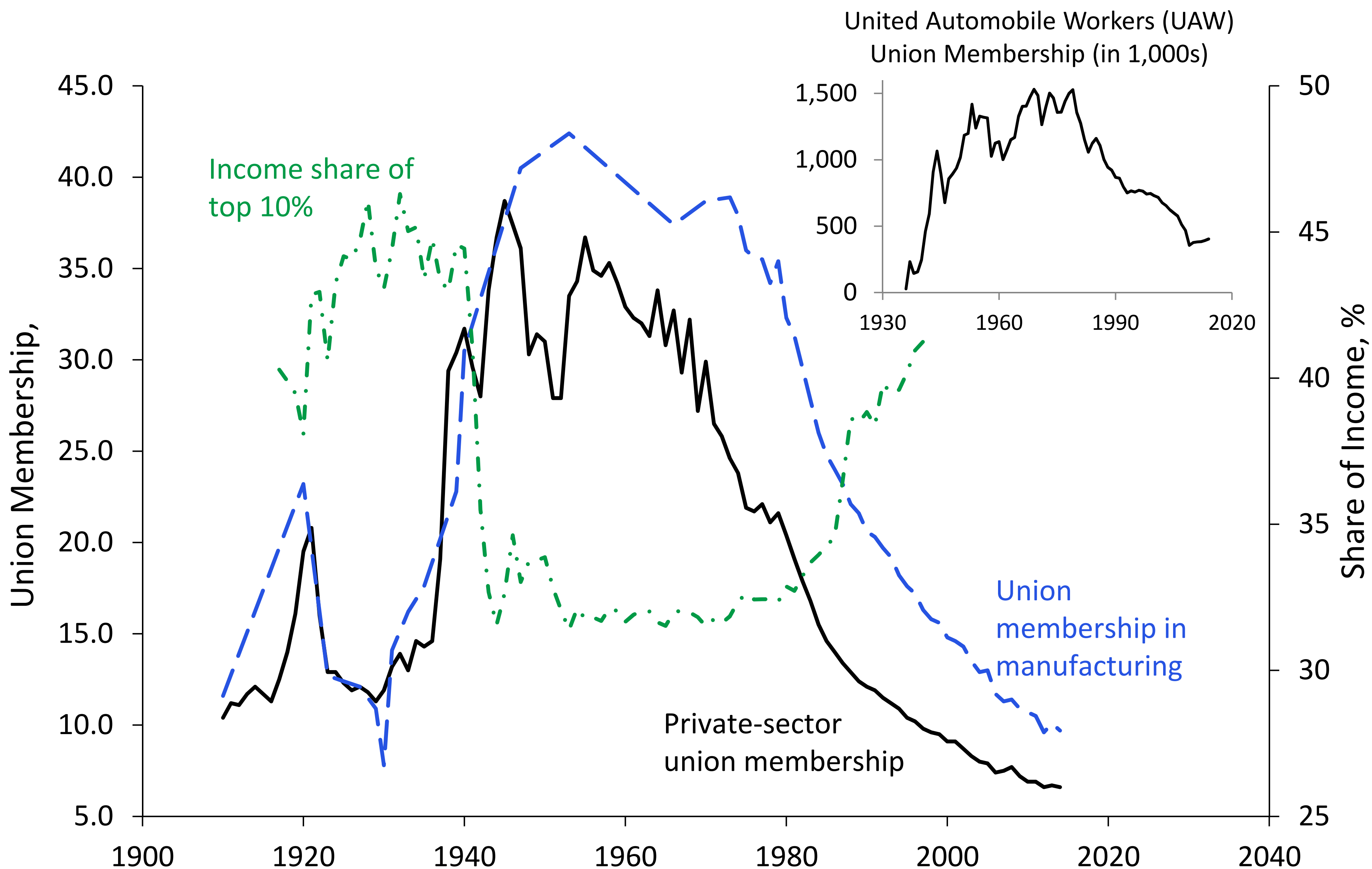


Figure 1. Union membership and income inequality in the United States. Notes: Private-sector and manufacturing union membership is from the Union Membership and Coverage Database and the U.S. Census Bureau. Income inequality is from Historical Statistics of the United States: Millennial Edition. UAW membership (inset) is from Walter P. Reuther Library archives. See the Online Appendix for more detail about the data and sources used in figures.

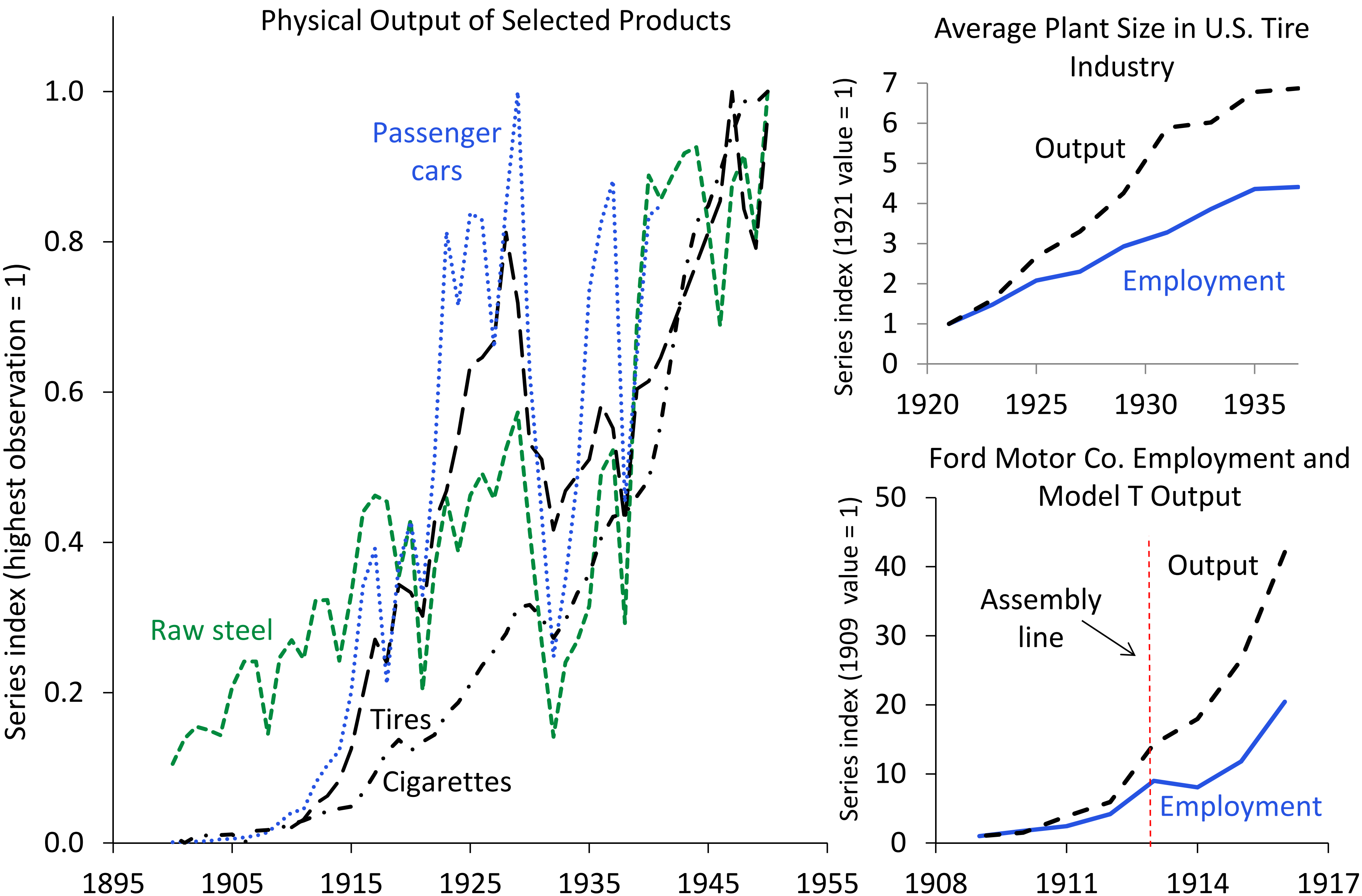


Figure 2. Physical output and plant size in selected U.S. manufacturing industries, 1900-1950. Notes: Physical output is from the Historical Abstract of the United States and the Census of Manufactures. Plant size in the U.S. tire industry is from French (1992). Ford Motor Co. employment and model-T output are from Williams, Haslam, and Williams (1992).

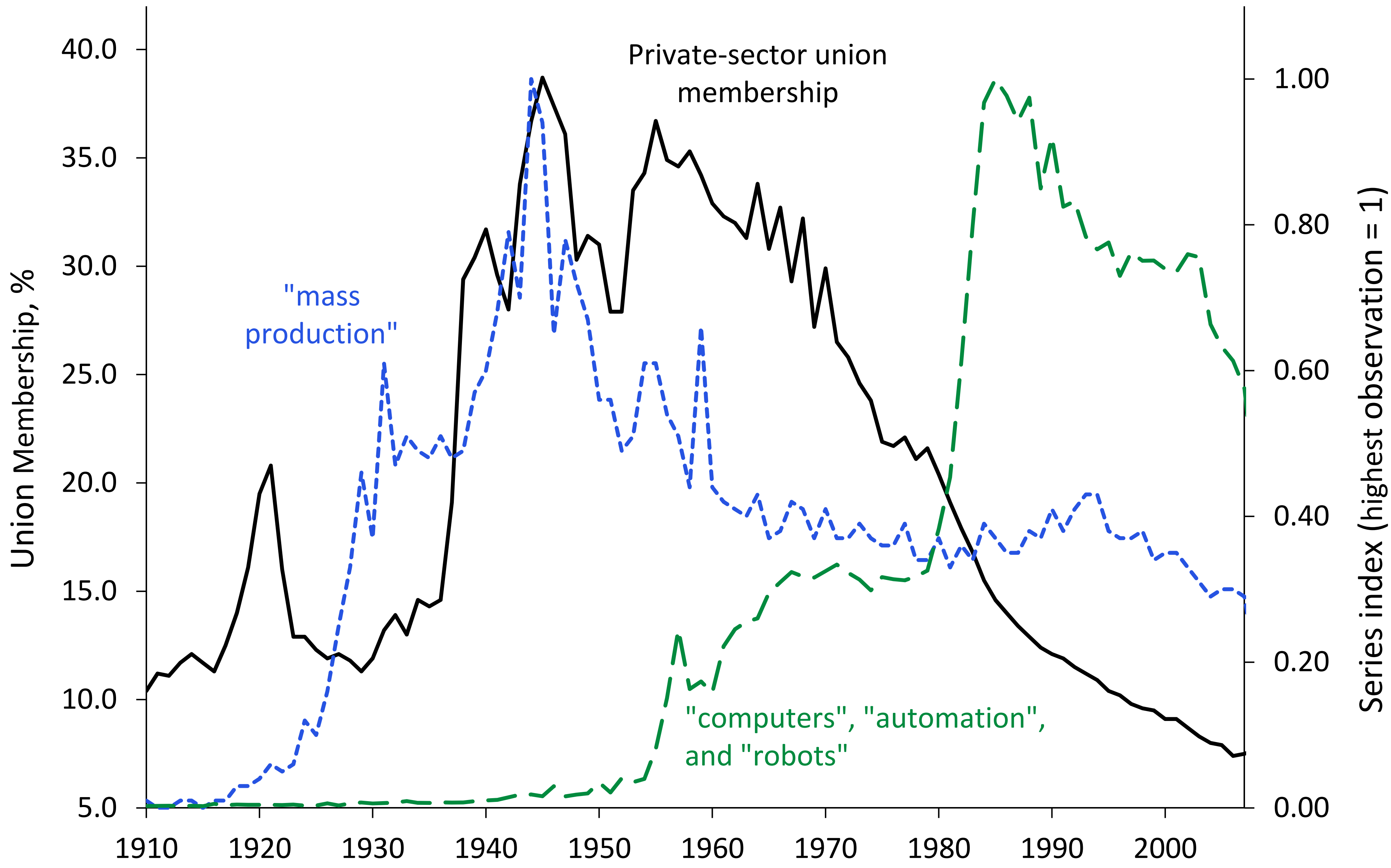


Figure 3. Union membership and technology in the United States. Notes: Private-sector union membership is from the U.S. Census Bureau; "mass production," "computers," "automation," and "robots" are based on the Google Ngram Viewer.

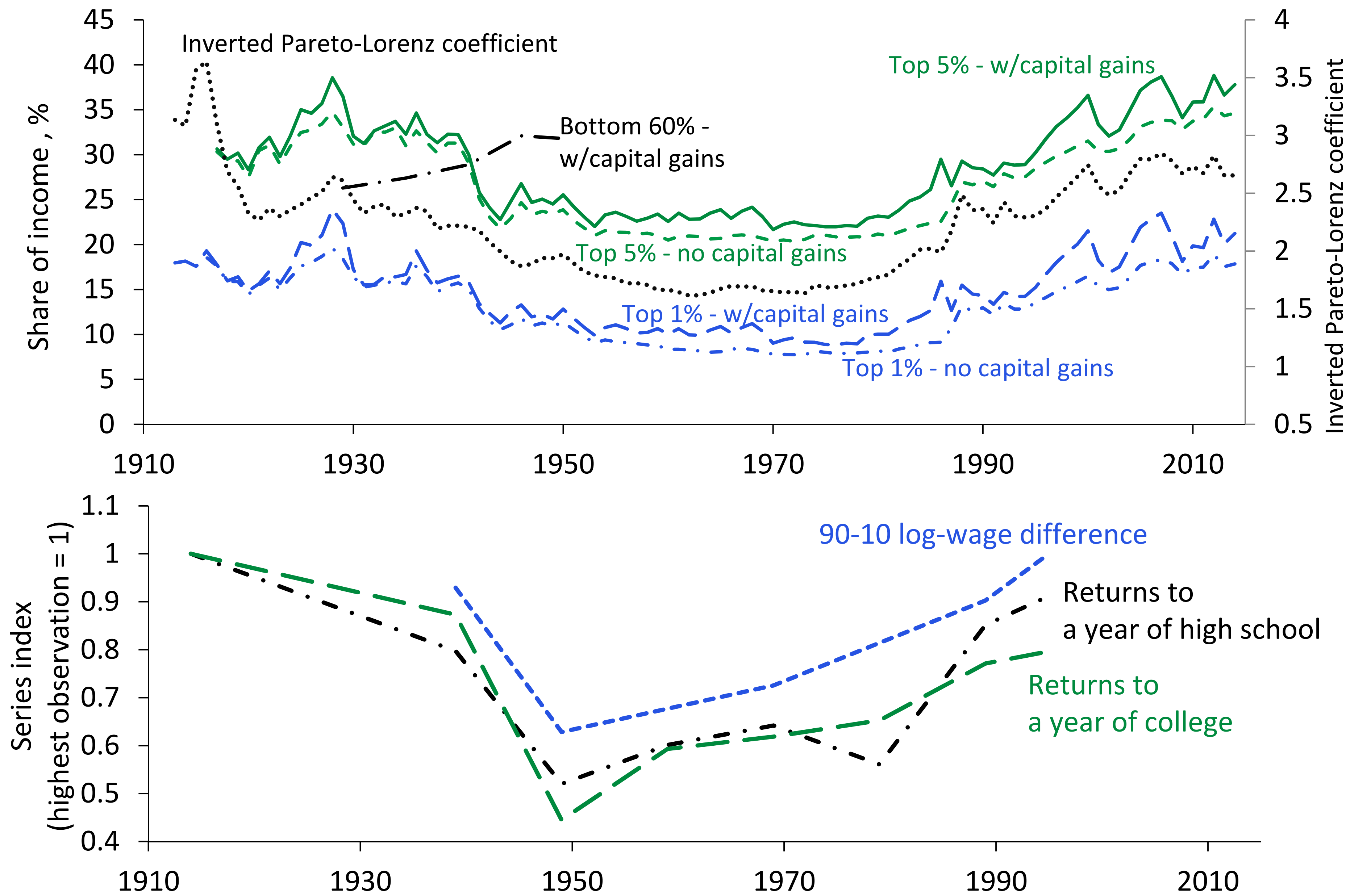


Figure 4. The evolution of income inequality, the wage gap, and the returns to schooling in the United States. Notes: The income inequality measures are from Alvaredo, Atkinson, Piketty, and Saez (2011) and McElvaine (2009). The 90-10 log-wage difference and the returns to high school and college are from Goldin and Katz (2001).

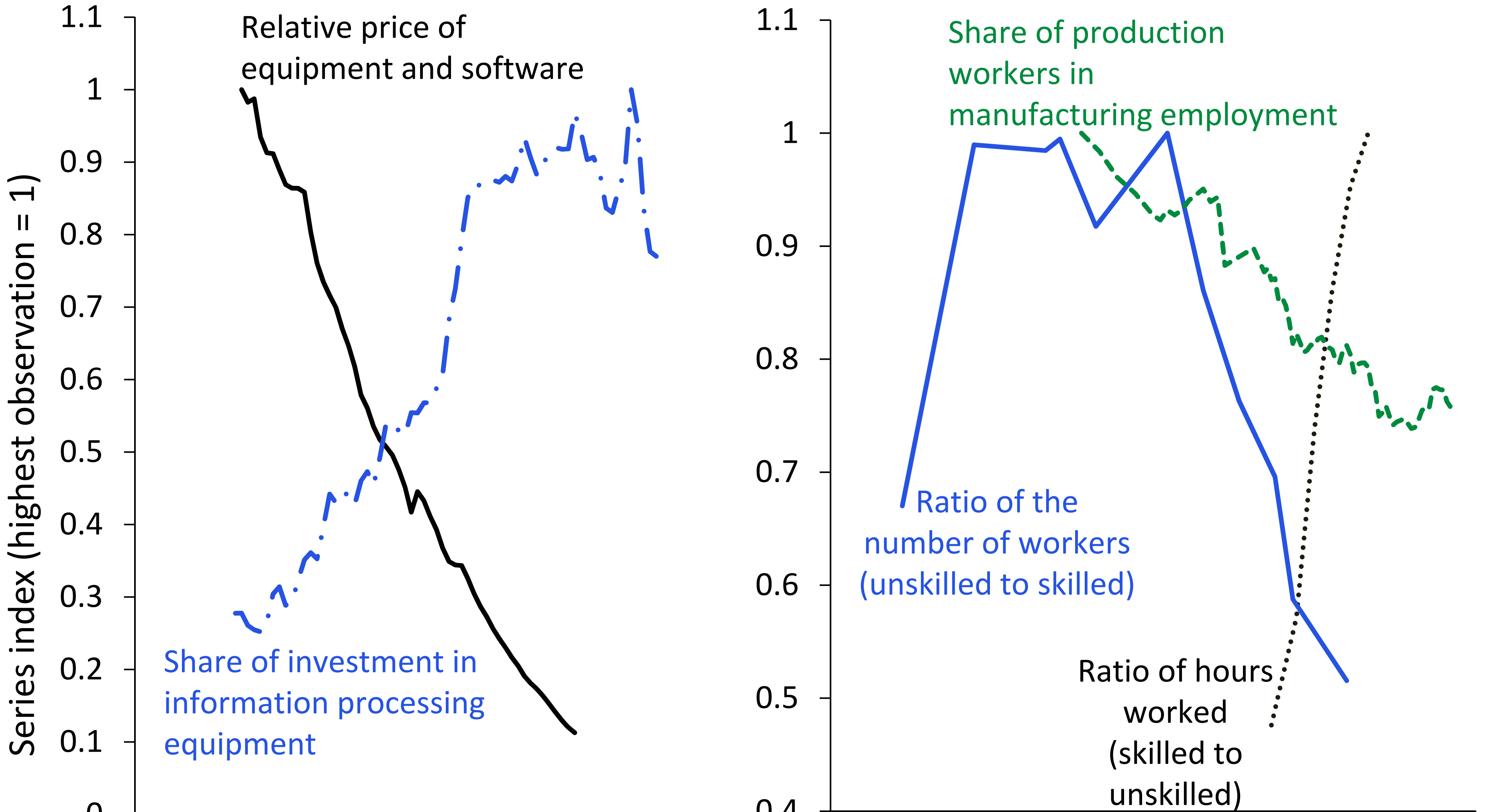


Figure 5. The evolution of the price of equipment and software and the skill composition in the United States. Notes: The relative price of equipment and software is from Cummins and Violante (2006). The share of investment in information processing equipment is from the Bureau of Economic Analysis. The ratio of hours worked is based on Krusell, Ohanian, Rios-Rull, and Violante (2000). The ratio of the number of workers is from the Historical Statistics of the United States: Millennial Edition (where unskilled labor is the sum of clerical workers, sales workers, operatives and laborers, and skilled labor consists of professionals, managers, officials and craft workers). The share of production workers is from the NBER-CES Manufacturing Database and the Census of Manufactures.

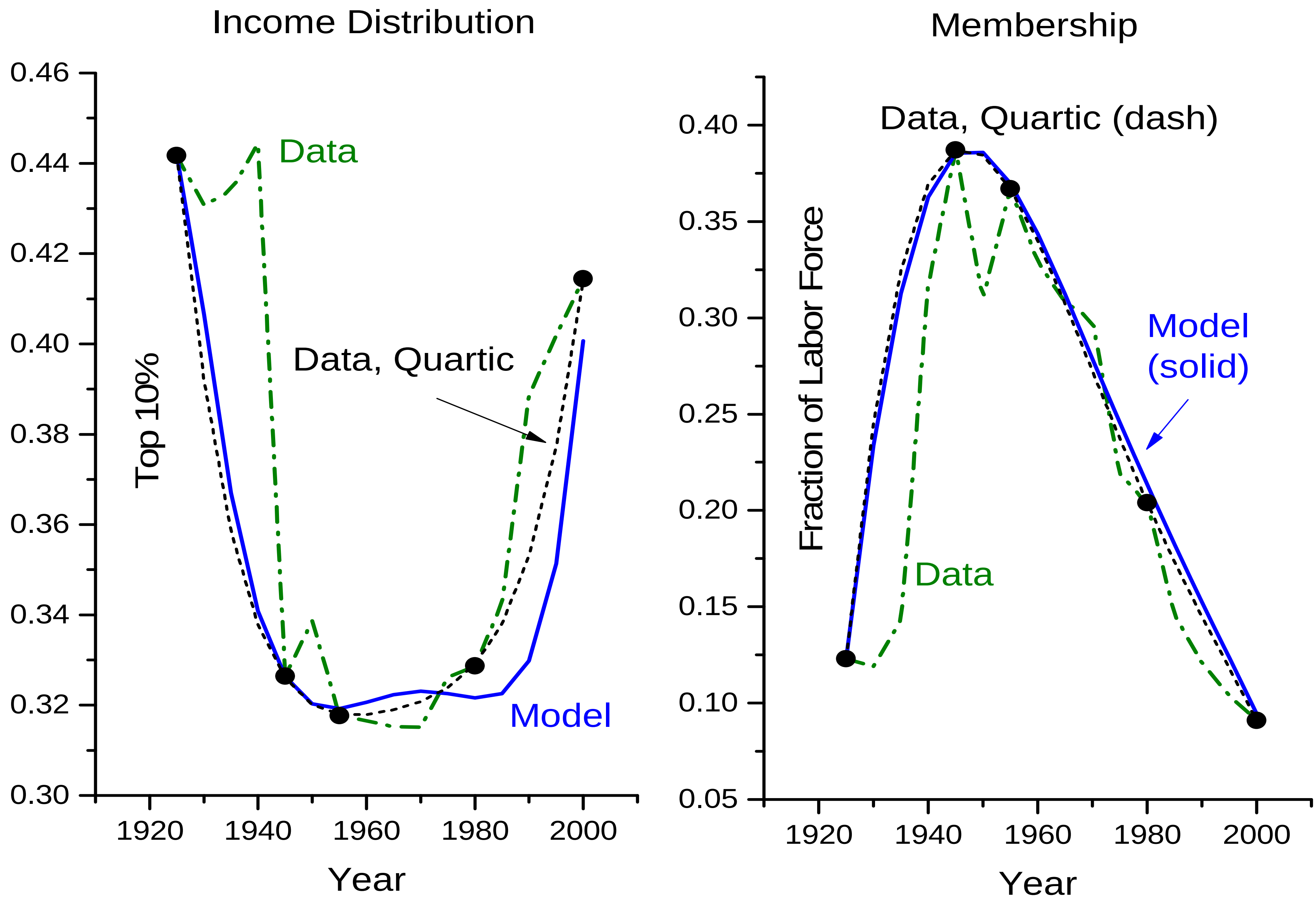


Figure 6. Union membership and income inequality--the data and the model. Notes: For the curves labelled "Data", private-sector union membership is taken from the Union Membership and Coverage Database and the U.S. Census Bureau, while income inequality is from Historical Statistics of the United States: Millennial Edition. The curves labelled "Data, Quartic" are the quartic polynomials fit to data. The curves labelled "Model" are based on the simulations of the model.

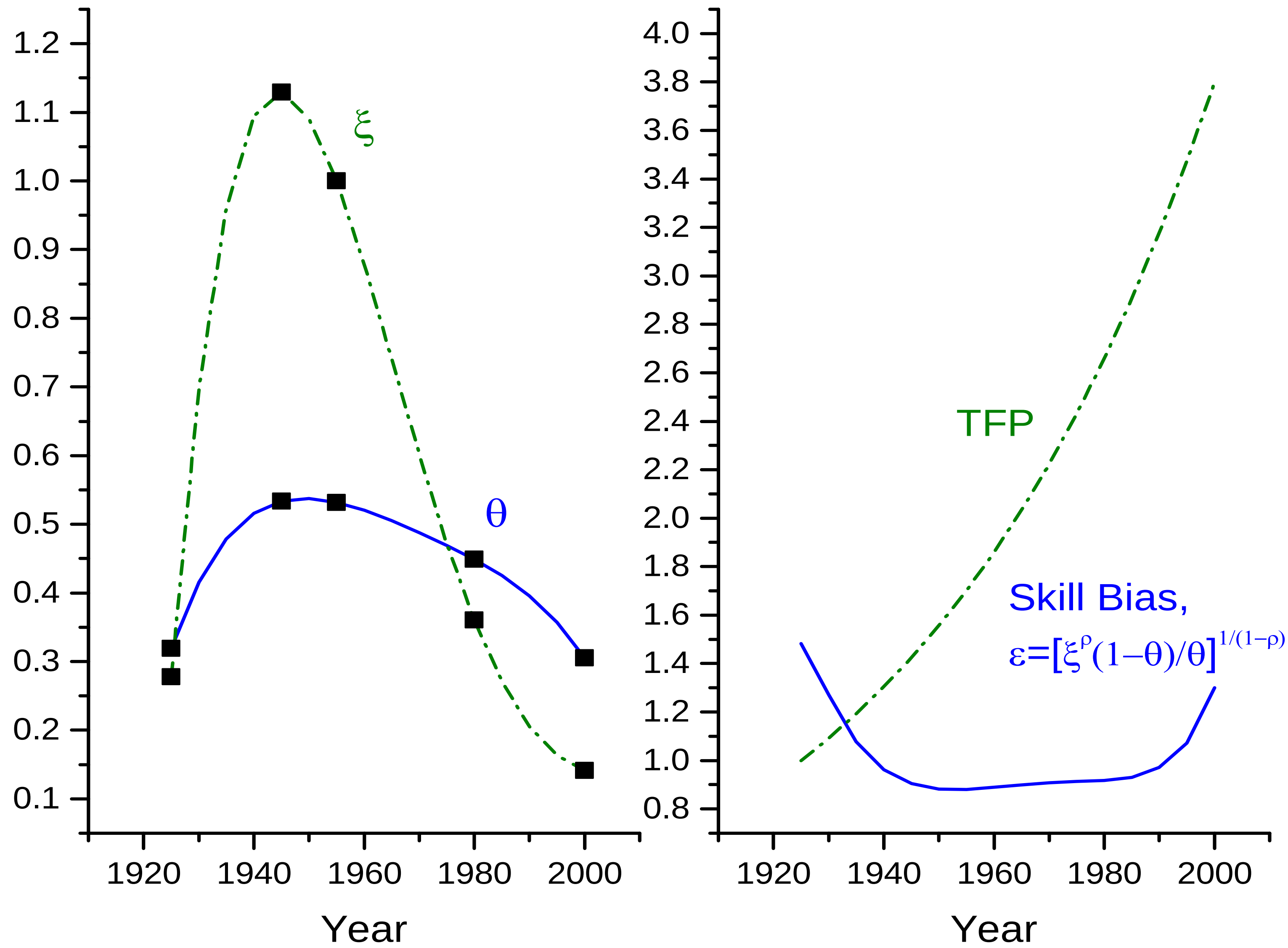


Figure 7. Skill-biased technological change in the model. Notes: The curves are based on the simulations of the model.

August 2016

Online Appendix

for

"The Rise and Fall of Unions in the United States"*

Emin Dinlersoz†

and

Jeremy Greenwood‡

Abstract

This document contains the Appendix accompanying "The Rise and Fall of Unions in the United States" by Emin Dinlersoz and Jeremy Greenwood.

---

*Any opinions and conclusions expressed herein are those of the authors and do not necessarily represent the views of the U.S. Census Bureau. All results have been reviewed to ensure that no confidential information is disclosed.

†Corresponding author. Center for Economic Studies, U.S. Census Bureau, 4600 Silver Hill Road, Suitland, MD 20746. Tel: (301) 763 7889, Fax: (301) 763 5935, E-mail: emin.m.dinlersoz@census.gov

‡Department of Economics, University of Pennsylvania, 3718 Locust Walk, Philadelphia, PA, 19104-6297.

Theory

Proof of Lemma 1 (Zero profits for the marginal firm). The proof is by contradiction. Suppose that the marginal firm does not earn zero profits or that constraint (9) is not binding. This implies that the constraint can be dropped from the optimization problem. Assume that an interior solution for unionization occurs. Then, the two first-order conditions associated with the union's problem will be

$$u_t : \omega[u_t - \frac{\chi_t m_t}{2} - w_t]^{\omega-1} m_t^{1-\omega}[1 - \frac{\chi_t}{2}\int_{z_t^u}^{\infty} \frac{dl_t^u(z)}{du_t} F(z)dz] + (1-\omega)[u_t - \frac{\chi_t m_t}{2} - w_t]^{\omega} m_t^{-\omega} \int_{z_t^u}^{\infty} \frac{dl_t^u(z)}{du_t} F(z)dz = 0,$$

and

$$z_t^u : (1-\omega)[u_t - \frac{\chi_t m_t}{2} - w_t]^{\omega} m_t^{-\omega} l^u(z_t^u)F(z_t^u) - \omega[u_t - \frac{\chi_t m_t}{2} - w_t]^{\omega-1} m_t^{1-\omega}\frac{\chi_t}{2} l^u(z_t^u)F(z_t^u) = 0.$$

[Recall that $l_t^u(z) = L_t^u(z_t^u; u_t, \cdot)$.] Take the second first-order condition and multiply it by $\int_{z_t^u}^{\infty}[dl_t^u(z)/du_t]F(z)dz$ to obtain

$$(1-\omega)[u_t - \frac{\chi_t m_t}{2} - w_t]^{\omega} m_t^{-\omega} \int_{z_t^u}^{\infty} [dl_t^u(z)/du_t]F(z)dz - \omega[u_t - \frac{\chi_t m_t}{2} - w_t]^{\omega-1} m_t^{1-\omega}\frac{\chi_t}{2} \int_{z_t^u}^{\infty} [dl_t^u(z)/du_t]F(z)dz = 0.$$

Using this in the first first-order condition then gives

$$\omega[u_t - \frac{\chi_t m_t}{2} - w_t]^{\omega-1} m_t^{1-\omega} = 0.$$

The last condition can only be true if

$$m_t = 0.$$

This cannot transpire, hence a contradiction. ■

Data Sources

Figure 1–Union membership and income inequality in the United States

Private-sector union membership. The private-sector union membership rates (excluding agriculture) for the period 1973-2012 come directly from the Union Membership and Coverage Database (unionstats.gsu.edu), with the year 1982 imputed using the average of the adjacent years. For the period 1910-1972, several versions of the U.S. Census Bureau's Statistical Abstract of the United States are used to calculate membership rates. Employment in the private sector (excluding agriculture) is used for the denominator of the union membership rate for all years 1910-1972. For the numerator, the number of union members excluding those in the public sector (federal, state, and local governments) is used. The number of total union members is available for all years 1910 to 1972. The numerator can be directly obtained from the available data items in the Abstract for the years 1910 to 1934. For the period 1935 to 1972, the number of public-sector union members is available for 1940, 1950, 1955, 1960, 1964-1966, 1968, and 1970. For these nine years, the share, h_t, of public-sector union members in total union members is as follows: 10% (1940), 13% (1950), 10% (1955), 12% (1960), 8% (1964), 9% (1965), 9% (1966), 11% (1968), and 13% (1970). For the rest of the years between 1940 to 1972 for which h_t is unavailable, a linear interpolation is used to impute a value, $\widehat{h}_t$. Private-sector union membership is then backed out as $1 - \widehat{h}_t$ times the total number of union members for each year for which h_t is unobserved. Alternative methods, such as using the grand average of h_t over the nine years for which it is available to impute the missing values, yield very similar time paths for the period 1935 to 1972. Note also that for the calibration of the model, the unionization rate is the one in 1955, which is not imputed.

Union membership in manufacturing. The manufacturing union membership rates for the period 1973-2012 are obtained from the Union Membership and Coverage Database, with the 1982 value imputed using the average of the adjacent years. For the period 1910-1972, the data comes from several resources. The data for the years 1930, 1935, 1939, 1940, 1947, 1953, 1966, and 1970 are from Table 3.63 in Troy and Sheflin (1985). The numbers for 1910, 1920, 1923, 1925, 1927, 1929, 1931, 1932 and 1933 are from Wolman (1937).

UAW union membership. The number of UAW union members is from the UAW records in the Walter P. Reuther Library of Labor and Urban Affairs at Wayne State University (http://reuther.wayne.edu/index.php). Only dues-paying members are included. William Lefevre, senior processing archivist, is thanked for providing the UAW membership figures.

Income share of top 10%. The underlying data for the income distribution comes from Historical Statistics of the United States: Millennial Edition. The data is series Be29. The measure used is the distribution of income among taxpaying units, specifically the share of income received by the upper 10%. Income is net of corporate taxes and employer-paid payroll taxes, but is before individual income taxes and individual-paid payroll taxes. It excludes capital gains. The series is based on work by Thomas Piketty and Emmanuel Saez. Similar series are available from Alvaredo, Atkinson, Piketty, and Saez (2011).

Figure 2– Physical output and plant size in selected U.S. manufacturing industries, 1900-1950

Physical output for selected industries. The data come from the Historical Abstract of the United States: Bicentennial Edition, Colonial Times to 1970.

Average plant size in the U.S. tire industry. The series (normalized by its 1921 value) is from Table 1 in French (1992).

Output and employment in the Ford Motor Company. The series (normalized by their 1909 values, respectively) are from Tables 1 and 2 in Williams, Haslam, and Williams (1992).

Figure 3–Union membership and technology in the United States

Private-sector union membership. Same as in Figure 1 above.

Mass production. The time series (normalized by its highest value over the sample period) is obtained by searching for the term "mass production" in the Google Ngram Viewer. The time series is the relative frequency of use of the term "mass production" in books published in English during the period 1910-2008, normalized by the number of books in each year–see the Google Ngram Viewer for a more detailed description of how this big data search engine works (https://books.google.com/ngrams/info). The series is very similar if one uses the terms "mass production" and "assembly line" together as search terms.

Computers, automation, and robots. The time series (normalized by its highest value over the sample period) is obtained by searching for the terms "computers," "automation," and "robots" in the Google Ngram Viewer. The series gives the total relative frequency of use associated with these terms in books published in English during the period 1910-2008. The series is very similar if one uses instead all of the terms "computers," "automation," "robots," "CNC" (Computer numerically controlled), "CAD" (Computer-aided design), "CAM" (Computer-aided manufacturing), "CIM" (Computer-integrated manufacturing), and "flexible manufacturing" together.

Figure 4–The evolution of income inequality, the wage gap, and the returns to schooling in the United States

Income inequality measures. The income shares of the top 1% and 5%, with or without capital gains, and the Inverted Pareto-Lorenz Coefficient are from Alvaredo, Atkinson, Piketty, and Saez (2011). The share of the bottom 60% is from the Table in page 331 of McElvaine (2009).

The 90-10 log wage difference, and the returns to high school and college. The series (normalized by their highest values, respectively) are from Figure 6 in Goldin and Katz (2001).

Figure 5–The evolution of the price of equipment and software and the skill composition in the United States

Relative price of equipment and software. The series (normalized by its 1947 value) is based on Cummins and Violante (2006).

The share of investment in information processing equipment. The series (normalized by its 1946 value) is investment in information processing and equipment as a fraction of total private investment as reported by the Bureau of Economic Analysis.

The ratio of hours worked (skilled to unskilled). The series (normalized by its 1990 value) is from Figure 3 in Krusell, Ohanian, Rios-Rull and Violante (2000).

The share of production workers in manufacturing employment. The series (normalized by its 1910 value) is from various years of the U.S. Census Bureau's Census of Manufactures, and all years of the NBER-CES Manufacturing Database.

The ratio of the number of workers (unskilled to skilled). The underlying data series come from the Historical Statistics of the United States: Millennial Edition. The unskilled labor force is taken to be the sum of clerical workers (Series Ba1038), sales workers (Ba1039), operatives (Ba1041), and laborers (Ba1045). The skilled workforce is professionals (Ba1034) plus managers and officials (Ba1037) added together with craft workers (Ba1040). In the figure the ratio of these two series is plotted.

Figure 6–Union membership and income inequality, the data and model

Private-sector union membership and income inequality (income share of the top 10%) values are from Figure 1.

Figure A1–Union membership versus the popularity of the term "labor unions"

Private-sector union membership (normalized by its highest value) is from Figure 1. The series "labor unions" (normalized by its highest value) is the relative frequency of use for the phrase "labor unions," obtained by searching for the term "labor unions" in Google Ngram Viewer for the period 1910-2008. The shape of the series is very similar if one also adds the phrase "trade unions."

Figure A2–Alternative proxies for income inequality in the United States

The income share of the top 10% is from Figure 1. All other series (normalized by their 1910 values) are based on the relative frequencies of use for the phrases "rich people," "middle class," "poor people," "high wage earner," "low wage earner," and "very rich people" obtained from the Google Ngram Viewer for the period 1910-2008. For example, the ratio

$$\frac{\text{“rich people”}}{\text{“rich people”} + \text{“middle class”} + \text{“poor people”}}$$

is the relative frequency of the phrase "rich people" with respect to the phrases "rich people," "middle class," and "poor people" combined.

Figure A3–Welfare cost of unions, Reesian analysis

Figure A3 illustrates the welfare cost of unions in the more traditional Reesian fashion. It draws the demands for unskilled labor by both union and non-union firms. These demands must sum up to 0.9, the size of the unskilled labor force as a proportion of the total labor force. In the economy without unions, the union firms would hire unskilled labor amounting to 49.6% of the total labor force at the competitive wage rate w^c. Unions increase this wage to u. As a consequence, unionized firms cut their employment of unskilled labor from 49.6% of the total labor force to 36.7%. This leads to a welfare loss measured by the area *acde*. But, the labor displaced by union firms is picked up by non-union ones. The wage rate for non-union labor falls from w^c to w. The gain in welfare from the increased employment by non-union firms is represented by the area *cdeb*. The net loss is the area in the triangle *acb*. This triangle represents the difference in productivities between the unionized and non-unionized firms. It amounts to $0.5 \times 0.20 \times w \times (0.496 - 0.367)$. Expressing this as a percentage of aggregate output, $\mathbf{o}$, gives

$$100\% \times 0.5 \times \underbrace{\frac{0.20 \times w \times 0.367}{\mathbf{o}}}_{0.0308} \times \underbrace{\frac{(0.496 - 0.367)}{0.367}}_{0.351} = 0.54\%.$$

This number is very close to the model-based figure of 0.66%. It is easy to see why this number is small. First, the union premium, 0.20, only applies to small part of wage bill expressed as a fraction of output, $w \times 0.367/\mathbf{o}$. This represents the base of the triangle.

Second, the proportional shift in union labor, $(0.496 - 0.367)/0.367 \simeq 0.351$, is not that large. This is the height of the triangle.

The Establishment Size Distribution

The statistics for the employee size of U.S. establishments used in the calibration of the model come from the Longitudinal Business Database (LBD) of the U.S. Census Bureau. The LBD contains annual information on employment for all U.S. establishments. For each year in the period 1976 to 2011, the mean, standard deviation, and the coefficient of variation of establishment employment are calculated across all establishments in the employer universe with at least one employee. The averages of the mean, standard deviation, and the coefficient of variation of employment over this period are approximately 17, 113, and 7, respectively. The average value (of about 7) for the coefficient of variation is used in guiding the calibration. The coefficient of variation stayed within a narrow band with a slight decline during this period. The coefficient of variation is slightly higher if establishments reporting zero employees are also included in the size distribution. Note also that comprehensive data for earlier years is difficult to find for this statistic. The earliest available data is for all manufacturing establishments in the 1963 Census of Manufactures, which reveals a CV of approximately 6 for employment. This value falls in the range of values calculated for the period 1976-2011.

References

Alvaredo, F., Atkinson, T., Piketty, T., Saez, E., 2011. The World Top Incomes Database. Available at http://topincomes.g-mond.parisschoolofeconomics.eu/.

Cummins, J., Violante, G., 2002. Investment-Specific Technical Change in the United States (1947-2000): Measurement and Macroeconomic Consequences. Review of Economic Dynamics 5, 243-284.

French, M., 1992. Structural Change and Competition in the United States Tire Industry, 1920-1937. The Business History Review, 60: 28-54.

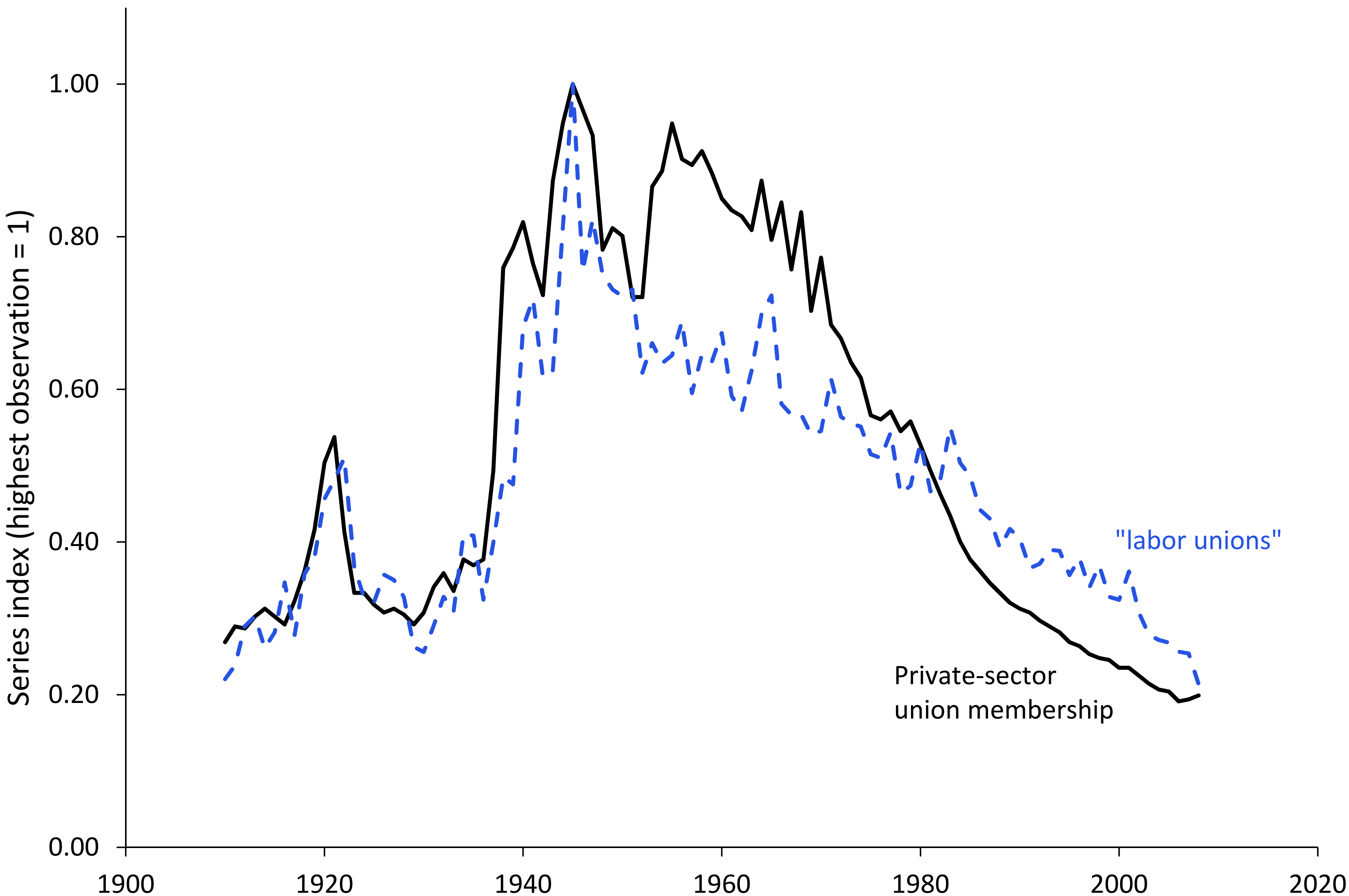
Krusell, P., Ohanian, L., Ríos-Rull, J., Violante, G., 2000. Capital-skill Complementarity and Inequality: A Macroeconomic Analysis. Econometrica 68, 1029-1053.

McElvaine, R., 2009. The Great Depression. New York: Three Rivers Press.

Troy, L., Sheflin, N., 1985. Union Sourcebook: Membership, Structure, Finance, Directory. First Edition. Industrial Relations Data Information Services.

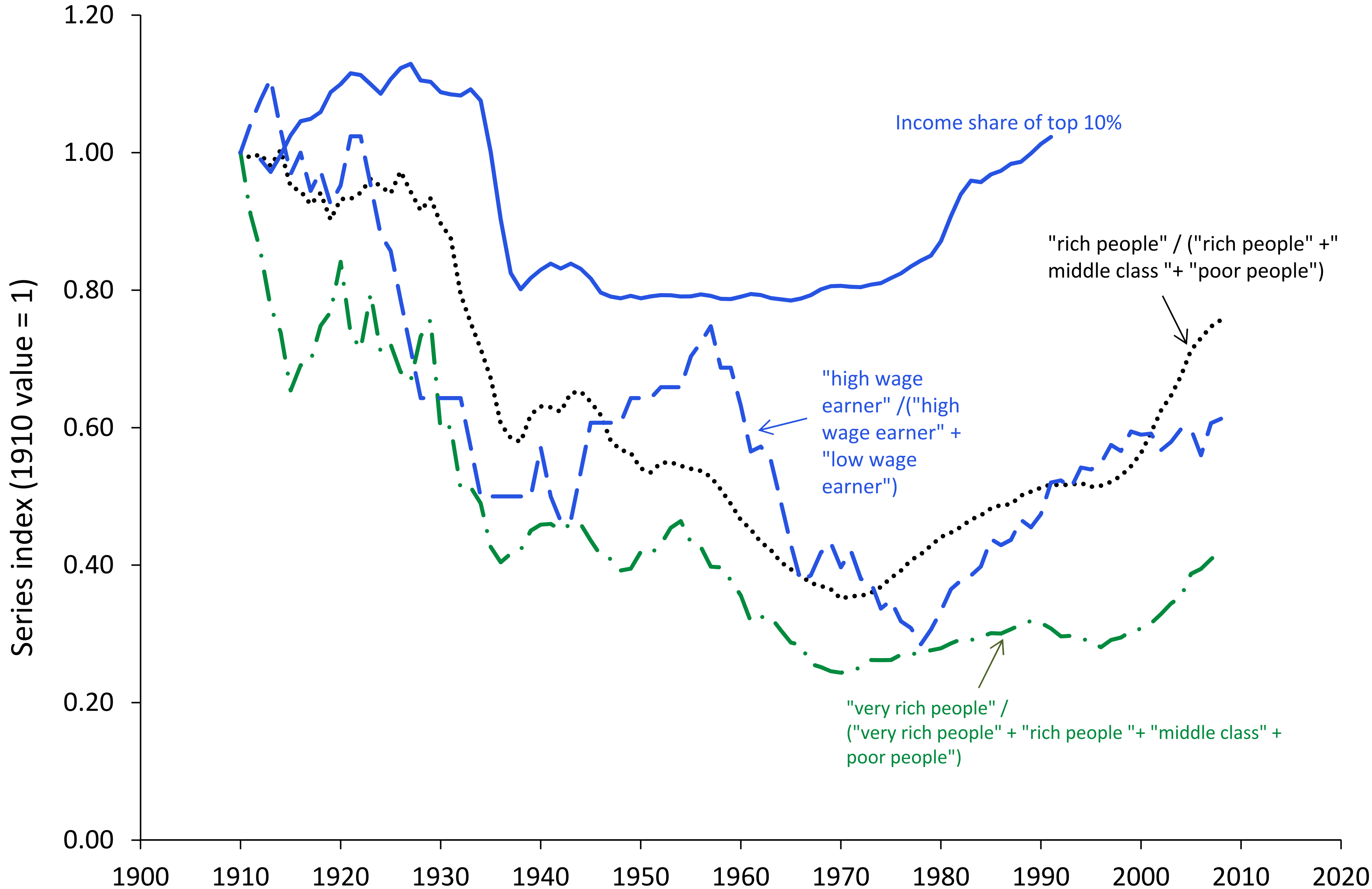
Williams, K., Haslam, C., Williams , J., 1992. Ford versus "Fordism": The Beginning of Mass Production? Work, Employment, and Society 6: 517-555.

Wolman, L., 1937. Union Membership in Great Britain and the United States. NBER Bulletin 68, December.



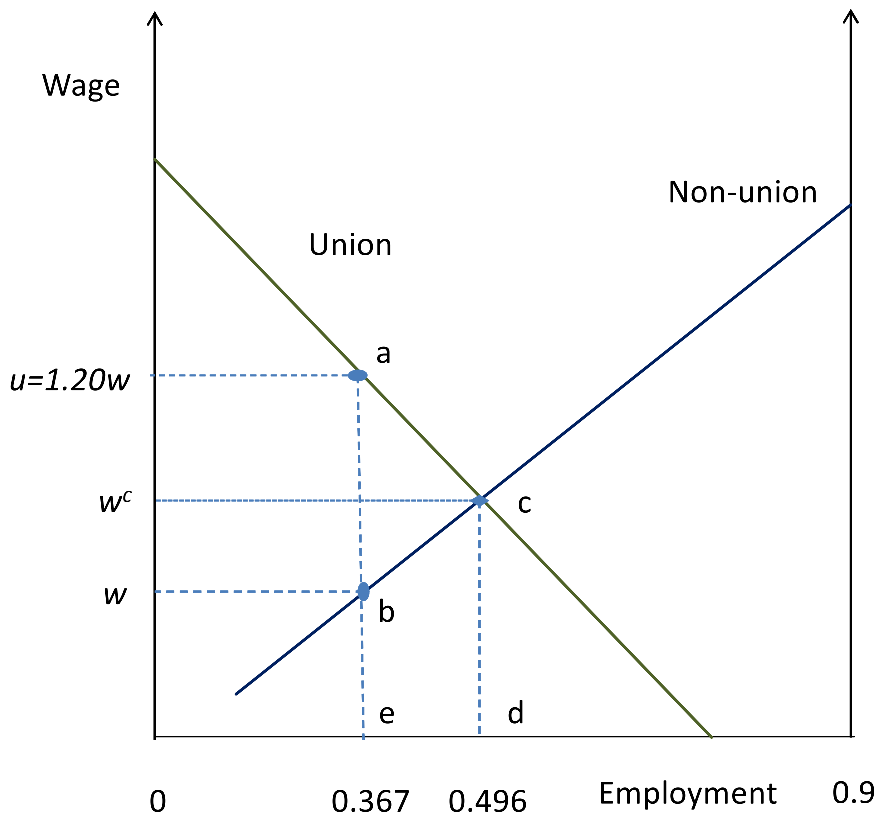
Source: Private-sector union membership – Union Membership and Coverage Database (http://unionstats.gsu.edu/) and U.S. Census Bureau; "labor unions" – Google Ngram Viewer (https://books.google.com/ngrams).

Figure A1. Union membership versus the popularity of the term "labor unions"



Source: Income inequality – Alvaredo, F., Atkinson, T., Piketty, T., Saez, E., 2011. The World Top Incomes Database (http://topincomes.g-mond.parisschoolofeconomics.eu/); “rich people”, “middle class”, “poor people”, “high wage earner”, “low wage earner”, “very rich people” – Google Ngram Viewer (https://books.google.com/ngrams).

Figure A2. Alternative proxies for income inequality in the United States



Source: Authors' calculations based on the model.

Figure A3. Welfare cost of unions – Reesian analysis