The Up and Down of Unions

by

Jeremy Greenwood

Abstract

Union membership displayed a \( \cap \)-shaped pattern over the 20th century, while the distribution of income sketched a \( \cup \). A model of unions is developed to analyze this phenomena. There is a distribution of firms in economy. Firms hire capital, plus skilled and unskilled labor. Unionization is a costly process. A union decides how many firms to organize and its members’ wage rate. The extent of unionization and the union wage rate is governed by the economy’s technology, specifically the productivity of skilled versus unskilled labor. Quantitative analysis illustrates how skilled-biased technological change, connected with mass production and computerization, can explain the above phenomena.

Keywords: Computerization; Distribution of Income; Mass Production; Skill-Biased Technological Change; Union Membership

Affiliation: University of Pennsylvania
1 Introduction

In 1900 seven percent of the American workforce were union members. The number of union members rose until the middle of the century, as shown in Figure 1, hitting its apex at 32 percent. It then began a slow decline. At the end of century 14 percent of American workers belonged to a union. At the beginning of the 20th century, the top 10 percent of workers earned 41 percent of income. This declined hitting a low of 31 percent around mid-century. It then steadily increased to 41 percent around 2000. What could have caused the \( \cap \)-shaped pattern of union membership and the \( \cup \)-shaped one for the distribution of income? Are they related?

The hypothesis here is that skilled-biased technological change underlies the up and down in union membership, along with the fall and rise in income inequality. The beginning of the 20th century witnessed a shift away from an artisan economy toward an assembly line one. This favored unskilled labor. The premium for skill declined. Unskilled labor is homogenous, almost by definition. This makes it easier unionize than skilled labor. When the demand for unskilled labor rises there is a larger payoff to unionizing it. Things changed 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. This represented a reversal of the earlier trend.

A general equilibrium model of unionization is developed. The union makes two interconnected decisions. First, it picks a common wage rate for its members. Second, the union selects which firms in the economy to organize. Unionization is a costly process. Firms sell output on 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 can pick a high wage. It also pays to organize more firms. Firms differ in their productivity, so when organizing labor the union will select the most profitable firms. Those firms that are not unionized can hire labor on competitive market.

The analysis builds upon the work of MacDonald and Robinson (1992). They present a
Figure 1: Union membership and the distribution of income over the 20th century
model of the extent of unionization in a competitive industry where all firms are the same. The key insights 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 things for an industry (or in partial equilibrium) and start off, in micro fashion, at the level of a firm’s cost function.

Modeling skill-biased technological change requires delving into a level deeper than the cost function; i.e., starting off from a firm’s production function. Analyzing the implications of this form of technological change for the economy’s distribution of income necessitates using a general equilibrium model. Furthermore, the current work investigates, in a quantitative sense, whether or the not such a framework is capable of explaining the extent of unionization and the level of income inequality that was observed over the course of the 20th century. It can.

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. Last, their analysis is theoretical in nature.

2 Mass Production and Computerization

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. This was achieved through the use of standardized parts, pioneered in the 19th century arms industry. Time spent fitting inexact parts was eliminated. The moving assembly line
was also inspired by the flow production techniques used in flour milling and meat packing. Greater specialization of labor was the result. It reduced the unnecessary handling of the product associated with ferrying the work between production operations—in early factories the placement of machines was often organized by their intrinsic operations (say drilling or milling) and not by where they lay in the production sequence.

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. They 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 work at the bench with other tools.” Labors were unskilled and did “manual labor that requires little or no experience or no judgement, such as shoveler, 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 classed as second class machinists and grouped with machine tenders and machine hands.” Meyer (1981, pp 13-14) describes how Ford engines were put together before 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 con-
ditions, a Ford motor assembler needed almost a full day of work to complete a single engine.

Mass production involved the breakdown of 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 40 percent 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 Table 2 illustrates. Operators make up the majority of workers. These were deskilled specialists performing routine machine operations. Mechanics account for only a small portion of the workforce. The deskilling of the workforce is nicely related by Wolmack et al (1990, pg. 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.

Only a few minutes of training was required to teach someone to be an assembler. This
system of manufacturing rapidly diffused through the American economy. The pinnacle of the mass production era was 1955.

Table 1: Workers in Detroit Metal Industries, 1891

<table>
<thead>
<tr>
<th>Occupation</th>
<th>No.</th>
<th>Percent</th>
<th>Mean Weekly Income</th>
</tr>
</thead>
<tbody>
<tr>
<td>Foreman</td>
<td>9</td>
<td>2</td>
<td>$19.67</td>
</tr>
<tr>
<td>Mechanics</td>
<td>153</td>
<td>39</td>
<td>12.58</td>
</tr>
<tr>
<td>Specialists</td>
<td>117</td>
<td>30</td>
<td>8.18</td>
</tr>
<tr>
<td>Unskilled Labor</td>
<td>113</td>
<td>29</td>
<td>6.60</td>
</tr>
<tr>
<td>Total</td>
<td>392</td>
<td>100</td>
<td>9.55</td>
</tr>
</tbody>
</table>

Source: Meyer (1981, pg. 46)

Table 2: Workers in Ford Motor Company, 1913

<table>
<thead>
<tr>
<th>Occupation</th>
<th>No.</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mechanics and Subforeman</td>
<td>329</td>
<td>2</td>
</tr>
<tr>
<td>Skilled Operators</td>
<td>3,431</td>
<td>26</td>
</tr>
<tr>
<td>Operators</td>
<td>6,749</td>
<td>51</td>
</tr>
<tr>
<td>Unskilled Workers</td>
<td>2,795</td>
<td>21</td>
</tr>
<tr>
<td>Total</td>
<td>13,304</td>
<td>100</td>
</tr>
</tbody>
</table>

Source: Meyer (1981, pg. 50)

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. The MIT machine would not have been reliable for commercial production; it had 250 vacuum tubes, 175 relays, and numerous moving parts. Programming them was a time consuming task. Standardized
languages had been developed for programming automated machine tools by the 1960s. At the same time the arrival of less expensive computers in the 1960s and 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 (CAD). Computers could also be used for planning and managing business in addition to running the machines on the factory floor (computer-aided manufacturing or CAM). In fact, sometimes they could automate virtually the entire business (computer-integrated manufacturing or CIM). 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 procedure 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 (or the forms) 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 production process. Quality control was at the end of the process in the typical mass production facility. Overtime Toyota gradually evolved to a process 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 part of the line. According to Wolmack et al (1990) in a mass production automobile plant about 20 percent of its area and 25 percent of working time are devoted to fixing mistakes. This is eliminated in a Toyota “lean production” facility. The Toyota production system favors skilled workers rather 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. This is shown in Figure 2, where unskilled workers are defined as clerical workers, laborers, operatives, and sales personnel, while skilled ones are taken to be craftsmen, managers, and professionals.

3 The Setting

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

\[ \sum_{t=1}^{\infty} \beta^{t-1} \ln c_t, \quad \text{with } 0 < \beta < 1, \]

where \( c_t \) 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 \( \sigma \) of these members are skilled, the rest unskilled. A skilled worker earns the period-\( t \) 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

Figure 2: Ratio unskilled to skilled workers, 1860 to 1990
receives $w_t$. The fraction of unskilled household members that work during period $t$ in the unionized part of the labor force is $p_t$, a variable that must be determined in equilibrium. The household saves in the form of physical capital. A unit of physical capital earns the rental $r_t$ in period $t$. Capital depreciates over time at the rate $\delta$. Finally, the household earns profits, $\pi_t$, from the firms that it owns. Firms are discussed now.

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

$$o_t = x_t z^\kappa k_t^\theta l_t^\rho + (1 - \theta_t) (\xi_t s_t)^\rho \alpha / \rho,$$

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 technological shift factor that is common across firms. An 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. In particular,

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

where $F$ is the density function for a Pareto distribution.

Observe that skilled and unskilled labor are aggregated via a CES production function. The technological variables $\theta_t$ and $\xi_t$ change over time and will capture the notion of skilled-biased technological change. Last, there is a fixed cost $\phi_t$ associated with operating a firm. There are diminishing returns to scale in production. All profits accruing from production are distributed to the household each period.

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

$$\frac{p_t \mu^\mu + 1}{\mu + 1},$$

where $p_t$ is the number of union members. These costs are recovered from the membership.
in the form of dues, \( d_t \). The union is given the following set of preferences:

\[
\sum_{t=1}^{\infty} \beta^{t-1} (u_t - d_t - w_t)^\omega p_t^{1-\omega}, \text{ with } 0 < \beta, \omega < 1.
\]

These preferences presume that the union has two regards. It values the surplus that a union member will earn over a non-unionized worker, \( u_t - d_t - w_t \), as well as the number of unionized workers, \( p_t \), that will receive it. As will be seen, there is a tradeoff involved with these two regards. A survey of the theory of unions is contained in Oswald (1985).

4 Decision Problems

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. This problem reads

\[
\max_{\{c_t, k_{t+1}\}_{t=1}^{\infty}} \sum_{t=1}^{\infty} \beta^{t-1} \ln c_t, \quad P(1)
\]

subject to

\[
c_t + k_{t+1} = (1 - \sigma) [(1 - p_t) w_t + p_t u_t] + \sigma v_t + (r_t + 1 - \delta) k_t + \pi_t \quad \text{for } t = 1, 2, \ldots .
\]

In the above maximization problem the household takes the number of union members, \( p_t \), as given. Given that \( u_t > w_t \), 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

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

for \( q = w, u \).  \( P(2) \)
With some abuse of notation, the variable \( q \) denotes if the firm is unionized \((q = u)\) or not \((q = w)\). 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 \( l^q_t(z) = L^q_t(z; q_t, \cdot) \)– 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 \( k^q_t(z) = K^q_t(z; q_t, \cdot) \) and \( s^q_t(z) = S^q_t(z; q_t, \cdot) \). The amount of output produced by a firm is denoted by \( o^q_t(z) = O^q_t(z; q_t; \cdot) \) and its profits are written as \( \pi^q_t(z) = \Pi^q_t(z; q_t, \cdot) \).

A firm will only produce if it makes nonnegative profits. Thus, it must transpire that in equilibrium

\[
\pi^q_t(z) = \Pi^q_t(z; q_t, \cdot) \geq 0, \text{ for } q = w, u.
\]

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

\[
\Pi^q_t(z^q_t; q_t, \cdot) = 0, \text{ for } q = w, u. \tag{1}
\]

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

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

\[
\frac{s^q_t}{l^q_t} = \left[ \xi^q_t \frac{(1 - \theta_t)}{\theta_t} \right] \times \frac{q_t}{v_t} \frac{1}{(1 - \rho)}.
\]

The ratio of skilled to unskilled labor, \( s^q_t/l^q_t \), in a firm depends on the price of unskilled labor relative to skilled labor, \( q_t/v_t \). It also depends on the technology term \( \xi^q_t \frac{(1 - \theta_t)}{\theta_t} \). This term will capture the notion of skilled-biased technological change in the model. When \( \xi^q_t \frac{(1 - \theta_t)}{\theta_t} \) is low, either because \( \xi_t \) is small or \( \theta_t \) is high, unskilled labor is favored, relatively speaking. The benefit of unionizing unskilled workers will be large. This is portrayed in Figure 3, where the slope of an isoquant is given by \( \frac{\theta_t}{(1 - \theta_t)\xi^q_t} \left[ (s^q_t/l^q_t)^{1 - \rho} \right] \). Thus, an upward shift in \( \theta_t \) or a rise in \( \xi_t \), causes the slope of an isoquant to increase along a ray from the origin. This is shown by the shift in the isoquant from \( S \) to \( U \). As a consequence, unskilled labor becomes more favored relatively speaking.
Figure 3: Skilled-biased technological change: unskilled labor becomes more favored when the isoquant shifts from $S$ to $U$
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. It is intuitive that the union should organize the firms with the highest level of productivity first. They can better afford to pay the union premium. There is 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 will not.

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 \( t \), \( p_t \), will be given by

\[
p_t = \int_{z^u_t}^{\infty} L_t^u(z; u_t, \cdot)F(z)dz.
\]

The dues paid by a union member, \( d_t \), are

\[
d_t = \frac{\chi P_t^{\mu+1}}{(\mu + 1)p_t} = \frac{\chi [\int_{z^u_t}^{\infty} L_t^u(z; u_t, \cdot)F(z)dz]^\mu}{1 + \mu}.
\]

The union’s decision problem appears as

\[
\max_{\{u_t, z^u_t\}_{t=1}^{\infty}} \sum_{t=1}^{\infty} \left\{ u_t - \frac{\chi [\int_{z^u_t}^{\infty} L_t^u(z; u_t, \cdot)F(z)dz]^\mu}{1 + \mu} - w_t \right\}^\omega [\int_{z^u_t}^{\infty} L_t^u(z; u_t, \cdot)F(z)dz]^{1-\omega}, \quad P(3)
\]

subject to the zero-profit constraint (1) holding (when \( q = u \)) for the marginal union firm, \( z^u_t \). When solving this problem, the union takes the wages for non-unionized unskilled and skilled labor, \( w_t \) and \( v_t \), as given. Is it possible that the union won’t pick the wage rate so that the threshold firm earns zero profits? The answer is no. Suppose that the marginal firm earned positive profits. The cost of raising the union wage incrementally is the loss of membership that will occur from all of the inframarginal firms. It turns out, though, that it this can be made up for by increasing the number of unionized firms or lowering \( z^u_t \). Therefore, this cannot be an optimal unionization strategy.

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

The union’s two regards must be traded off in the above maximization problem. By applying the envelope theorem to a unionized firm’s optimization problem \( P(2) \) for \( q = u \) it can be easily calculated that

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

This implies that lowering the threshold hold, \( z_t^u \), or equivalently unionizing more firms, can only accomplished by reducing the union wage, \( u_t \). Additionally, it can be seen in the objective function that a rise in membership, \( p_t = \int_{z_t^u}^{\infty} L_t^u(z; u_t, \cdot) F(z) dz \), 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

\[
\int_{z_t^u}^{\infty} k_t^w(z) F(z) dz + \int_{z_t^u}^{\infty} k_t^u(z) F(z) dz = k_t. \tag{2}
\]

The market-clearing condition for skilled labor is

\[
\int_{z_t^u}^{\infty} s_t^w(z) F(z) dz + \int_{z_t^u}^{\infty} s_t^u(z) F(z) dz = \sigma, \tag{3}
\]

while that for unskilled labor reads

\[
\int_{z_t^u}^{\infty} l_t^w(z) F(z) dz + \int_{z_t^u}^{\infty} l_t^u(z) F(z) dz = 1 - \sigma. \tag{4}
\]

Last, equilibrium in the goods market implies

\[
c_t + k_{t+1} = \int_{z_t^p}^{z_t^u} o_t^w(z) F(z) dz + \int_{z_t^u}^{\infty} o_t^u(z) F(z) dz + (1 - \delta) k_{t+1}. \tag{5}
\]

A definition of the equilibrium under study will now be presented to take stock of the situation so far.
Definition 2 A competitive equilibrium is a time path for consumption and savings, \( \{c_t, k_{t+1}\}_{t=1}^{\infty} \), a set of labor and capital allocations for union \((q = u)\) and non-union \((q = w)\) firms \(\{l_t^u(z), s_t^u(z), k_t^u(z)\}_{t=1}^{\infty}\), a set of factor prices \(\{u_t, w_t, v_t, r_t\}_{t=1}^{\infty}\), and a sequence determining the threshold points for non-union and union firms, \(\{z_t^w, z_t^u\}_{t=1}^{\infty}\), such that for a given time profile for technology \(\{\theta_t, \xi_t, x_t\}_{t=1}^{\infty}\):

1. The time path for consumption and savings, \(\{c_t, k_{t+1}\}_{t=1}^{\infty}\), solves the representative household’s problem, \(P(1)\), given the time path for factor prices, \(\{u_t, w_t, v_t, r_t\}_{t=1}^{\infty}\), profits, \(\pi_t = \int_{z_u^t}^{z_w^t} \pi_t^u(z)F(z)dz + \int_{z_u^t}^{\infty} \pi_t^w(z)F(z)dz\), and the size of the union sector, \(\rho_t = \int_{z_u^t}^{z_w^t} l_t^u(z)F(z)dz\).

2. The time paths for firms’ input utilizatons, \(\{l_t^u(z), s_t^u(z), k_t^u(z)\}_{t=1}^{\infty}\), solve their profit maximization problems, as specified by \(P(2)\), given the time paths for factor prices, \(\{q_t, v_t, r_t\}_{t=1}^{\infty}\) (for \(q = u, w\)).

3. The sequence for union wages, \(\{u_t\}_{t=1}^{\infty}\), and the threshold, \(\{z_t^u\}_{t=1}^{\infty}\), solve the union’s problem \(P(3)\), given the time paths for competitive wages, \(\{w_t, v_t\}_{t=1}^{\infty}\), the rental rate for capital, \(\{r_t\}_{t=1}^{\infty}\), 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 \(P(2)\).

4. The sequence for non-union thresholds, \(\{z_t^w\}_{t=1}^{\infty}\), solves (1) when \(q = w\), given \(\pi_t^w(z) = \Pi_t^w(z; w_t, \cdot)\) from \(P(2)\).

5. The markets for capital, labor and goods, all clear so that (2) to (5) hold.

6 Calibration

Before the model can be simulated values must be assigned for its parameters. The period is taken to be five years. Given this, the discount factor is set so \(\beta = 1/(1.04)^5\), which implies an annual interest rate of 4 percent. This is a standard value. The depreciation rate
for capital is taken to be 0.08, another standard value. Likewise, labor’s share of income is set at 60 percent, implying $\alpha = 0.60$, another typical value if one assumes that part of the capital stock includes intangibles. Note that a firm’s production function exhibits diminishing returns to scale. Guner, Ventura and Xi (2008) estimate that the share of profits in output is 20 percent. Given this, capital’s share of income, $\kappa$, is set at 0.20. Katz and Murphy (1992) estimate that the elasticity of substitution between skilled and unskilled labor is 1.4. This corresponds to a value of 0.29 for $\rho$.

The rest of the model’s parameters are selected so that a steady state for the model hits 5 data targets for the year 1955. This is the peak of the unionization movement. This involves computing the model’s steady state in conjunction with the 5 data targets while taking the 5 parameters $\theta_{1955}$, $\mu$, $\omega$, $\chi$ and $\zeta$ as additional variables. The technological variable $\xi$ is normalized to one for 1955 so that $\xi_{1955} = 1$. The first target is the fraction of population that was unionized. In 1955 this was 32 percent. Therefore, the steady state is computed subject to restriction

$$p_{1955} = 0.32.$$  

Let the top 10 percent of the population represent skilled labor. Thus, $\sigma = 0.10$. The share of the top 10 percent of the work force in earning was 0.32. Therefore, the steady state must satisfy the condition

$$\frac{\sigma v}{pu + (1 - p - \sigma)w + \sigma v} = 0.32.$$  

Union dues are assumed to amount to 1 percent of a union members wages. MacDonald and Robinson (1992, pg 47) state that this is a reasonable value. Indeed, this is exactly what the UAW currently charges salaried workers. Thus, set

$$\frac{d_{1955}}{u_{1955}} = 0.01.$$  

The distribution of employment across establishments in the U.S. is very tight. Henley and Sanchez (2009, p 427) report that the coefficient of variation across U.S. establishments was 8 percent in 1974. It remained relatively constant after that. This observation is targeted to provided guidance for the choice of the Pareto distribution parameter $\zeta$. Next, the wage
premium from a union membership is taken to 0.15 percent, the famous value of H. Greg Lewis. This implies

\[
\frac{u_{1955}}{w_{1955}} = 0.15.
\]

### Table 1: Parameter Values

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Definition</th>
<th>Basis</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Tastes</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \beta = (1.03)^{-5} )</td>
<td>discount factor</td>
<td>standard</td>
</tr>
<tr>
<td><strong>Technology</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \alpha = 0.60 )</td>
<td>labor’s share</td>
<td>Greenwood et al (2010)</td>
</tr>
<tr>
<td>( \delta = 0.08 )</td>
<td>depreciation rate</td>
<td>standard</td>
</tr>
<tr>
<td>( \kappa = 0.20 )</td>
<td>exponent on capital</td>
<td>Guner et al (2008, p 732)</td>
</tr>
<tr>
<td>( \rho = 0.29 )</td>
<td>elasticity of substitution</td>
<td>Katz and Murphy (1992, eq 19)</td>
</tr>
<tr>
<td>( \theta_{1955} = 0.45 )</td>
<td>weight on skilled labor</td>
<td>data targets</td>
</tr>
<tr>
<td>( \xi_{1955} = 1.0 )</td>
<td>shift factor on skilled labor</td>
<td>normalization</td>
</tr>
<tr>
<td>( \zeta = 35 )</td>
<td>Pareto distribution</td>
<td>data target</td>
</tr>
<tr>
<td>( \phi = )</td>
<td>fixed cost</td>
<td>data target</td>
</tr>
<tr>
<td><strong>Unionization</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \omega = 0.65 )</td>
<td>ideals–wage</td>
<td>data targets</td>
</tr>
<tr>
<td>( \mu = 1.1 )</td>
<td>organization costs, exponent</td>
<td>data targets</td>
</tr>
<tr>
<td>( \chi = 0.03 )</td>
<td>organization costs, constant</td>
<td>data targets</td>
</tr>
</tbody>
</table>

### 7 Results

Can the model explain the \( \cap \)-shaped pattern of union membership along with the \( \cup \)-shaped profile for income inequality that were observed over the 20th century? To investigate this question requires inputting in a time series process for technology, \( \{\theta_t, \xi_t\}_{t=1920}^{2000} \). Such a process is constructed in a crude way. Steady states for the model are computed for 1920 and 2000, the starting and ending years for the analysis. Union membership and
income inequality are taken as targets for these years. Solutions for $\theta$ and $\xi$ are backed out that hit these targets, while holding all other parameter values fixed. Assume that $\theta_t$ and $\xi_t$ are separately quadratic in $t$. Each quadratic will have three parameters. Fit these two quadratics to the triplets $(\theta_{1920}, \theta_{1955}, \theta_{2000})$ and $(\xi_{1920}, \xi_{1955}, \xi_{2000})$, respectively. The resulting time profile skill-biased technological change, as represent by $[\xi_t^u(1 - \theta_t)/\theta_t]^{1/(1-\rho)}$, is shown in Figure 4. Is this pattern of skill-biased technological change reasonable? The extent of the required shift is quite modest, 25 percent from peak to trough. Over the 1920 to 2000 period real per-capita income grew by 2.25 percent a year. This implies that real per-capita GDP rose by a factor of 6. To achieve this in the model, the neutral technological shift factor, $x$, must rise by a factor of $6^{1/(1-\kappa)} = 9.4$. Therefore, skilled bias technological change is swamped by neutral technological change. Imagine calculating total factor productivity as is conventionally done by $\text{TFP}_t = [\int_{1}^{z_1} o_t^u(z) F(z) dz + \int_{z_1}^{\infty} o_t^u(z) F(z) dz]/k^x_t$. All the observer would see is a seemingly esmooth rise in TFP, as Figure 4 illustrates. He would not notice the tiny wiggles associated with skilled-biased technological change.

The framework does a good job accounting for the rise and fall in union membership, as Figure 5 illustrates. It also mimics the fall and rise in income equality as well. This is shown in Figure 6.

### 7.1 Welfare Cost of Unions

So, what is the welfare cost of unions? Rees (1963) asked this question a long time ago. He found that the welfare loss from unions in 1957 amounted to 0.14 percent of GDP. The model developed here can 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. Welfare would only increase by 0.38 percent of GDP. While this is 2.7 times as big as Rees’s number, it is paltry.

Figure 7 illustrates the situation in Reesian fashion. The picture 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
Figure 4: Technological Change, model
Figure 5: Union Membership over the 20th Century, data and model

Figure 6: The Distribution of Income over the 20th Century, data and model
economy without unions, the union firms would hire 0.42 percent of the total labor force (including skilled labor) 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 0.42 percent of the total labor force to 0.32 percent. This leads to 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 nonunion 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 $0.5 \times 15 \times w \times (0.42 - .32)$. Expressing this as a percentage of aggregate output, $o$, gives

$$
100\% \times \frac{0.5 \times .15 \times w \times (0.42 - .32)}{o} = 100\% \times 0.5 \times \frac{0.15 \times w \times 0.32}{0.021} \times \frac{(0.42 - 0.32)}{0.31} = 0.32\%.
$$

It is easy to see why this number will be small. First, the union premium, 0.15, only applies to small part of wage bill expressed as a fraction of output, $w \times 0.32/o$. This represents the base of triangle. Second, the proportional shift in union labor, $(0.42 - 0.32)/0.32$, is not that large. This is the height of the triangle. Note that this triangle estimate is extremely close to general equilibrium one. Rees’s (1963) is a bit vague on the details surrounding his estimation. First, it is unclear how he obtains the magnitude of his shift in labor demand. Second, it is hard to tell whether his estimated shift in labor should refer to the length $ed$ or the length $bf$. If it is the later, as it seems from the context he took the number from, then it is too large. Doing an appropriate correction would yield something around 0.07 percent of GDP. The difference between Rees’s estimate and the current one derives from a difference in the implied elasticities for labor demands.
Figure 7: The welfare loss from unions
8 Conclusions

A general equilibrium model of unionization is developed here. Firms hire capital, skilled labor and unskilled workers. They differ in their productivity. A union can organize unskilled labor, but at a cost. It cares about the wage rate that its members will earn. It also is concerned about how many workers will receive this wage. There is a trade off between these two objectives. The union sets the wage so that it squeezes all of the rents from the last firm organized. The higher is the union wage the smaller is the number of unionized of firms and the amount of unskilled labor that each will hire.

The structure of production influences that value of unskilled labor in economy. When the productivity of unskilled labor is (relatively) high it pays for the union to organize a lot of firms and demand overly generous wages. It is argued here that the shift from an artisan economy to an assembly line one 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 the assembly line economy and the rise of the information age during the second half of the century reversed this. This led to the \( \cap \)-shaped pattern of unionization and the \( \cup \)-shaped one for income inequality.

9 Appendix

9.1 Data

Figure 1. The data is taken from the Historical Statistics of the United States: Millennial Edition. Union membership is taken from three series: Series Ba4789 for 1890 to 1914; Series Ba4783 for 1915 to 1976; Series Ba4788 is for 1977 to 1999. The union membership series is then divided through by a measure of the labor force. For this, the total civilian labor force is taken from Series Ba471. The farm labor force is netted out of this series. For 1890 to 1990, Series Ba472 is used for the farm labor force. Series Ba482 gives the data for 1991 to 1999. The data on income distribution is series Be29 and refers to the distribution of income
among taxpaying units, specifically the share of income received by the 10th percentile.

Figure 2. The underlying data series come from *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.

9.2 Theory

Proof. – Lemma 1. Suppose not and that an interior solution for unionization occurs. Then, the two first-order conditions associated with the above problem will be

\[
\omega[u_t - \frac{\chi p_t^\mu}{1 + \mu} - w_t] = \omega^{-1} p_t^{1-\omega} [1 - \mu \frac{\chi p_t^\mu}{\mu + 1} \int_{z_t^u} d l_t^u(z) F(z) dz]
\]

\[
+ (1 - \omega)[u_t - \frac{\chi p_t^\mu}{\mu + 1} - w_t] p_t^{-\omega} \int_{z_t^u} d l_t^u(z) F(z) dz = 0,
\]

and

\[
(1 - \omega)[u_t - \frac{\chi p_t^\mu}{1 + \mu} - w_t] p_t^{-\omega} \int_{z_t^u} l_t^u(z) F(z) dz
\]

\[
\geq \omega[u_t - \frac{\chi p_t^\mu}{1 + \mu} - w_t]^{-1} p_t^{1-\omega} \int_{z_t^u} d l_t^u(z) F(z) dz.
\]

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

\[
(1 - \omega)[u_t - \frac{\chi(p_t)^\mu}{1 + \mu} - w_t]^{-\omega} p_t^{-\omega} \int_{z_t^u} [d l_t^u(z)/du_t] F(z) dz
\]

\[
- \omega[u_t - \frac{\chi(p_t)^\mu}{1 + \mu} - w_t]^{-1} p_t^{1-\omega} \int_{z_t^u} [d l_t^u(z)/du_t] F(z) dz = 0.
\]

Using this in the first first-order condition then gives

\[
\omega[u_t - \frac{\chi p_t^\mu}{1 + \mu} - w_t] = 0.
\]
The last condition can only be true if
\[ u_t - \frac{\chi(p_t)^\mu}{\mu + 1} - w_t = 0. \]
This cannot transpire, hence a contradiction. ■

References


