Sluggish Institutions in a Dynamic World: Can Unions and Industrial Competition Coexist?

Barry T. Hirsch

During the late 1930s, 1940s, and 1950s, collective bargaining emerged as the industrial workplace norm in the United States. As recently as 30 years ago, unionization was widespread among private sector production workers in U.S. manufacturing, construction, mining, transportation, communications, and utilities. But the remarkable emergence of industrial unionism was then followed by a no less remarkable, albeit gradual, union decline. In the U.S. private sector, union density has declined from one-in-four wage and salary workers being a union member in the early 1970s to under one-in-thirteen today. In manufacturing, density has fallen from two of every five workers in the early 1970s to under one-in-eight currently.\(^1\) This paper examines the shift away from union governance in the U.S. private sector toward the current norm of constrained employer discretion. Although declines in union density are seen in many countries (Blanchflower, 2006), discussion is focused on the United States because of substantial international differences in workplace institutions, laws, and norms.

A useful starting point is to ask what U.S. unions do in the workplace. At worksites where a majority of employees have voted for representation, a union serves as workers’ exclusive bargaining agent and provides a collective voice in dealings with employers. Union goals and policies are most influenced by the

---

\(^1\) Union density is defined as the percent of wage and salary workers who are union members. Unless otherwise stated, all union membership and density figures have been compiled from the Current Population Survey, either by myself for this paper or in collaboration with David Macpherson (Hirsch and Macpherson, 2003; database updated annually and posted at [http://www.unionstats.com]). A data appendix providing all values shown in Figures 1–5 is posted at the Unionstats database.

Barry T. Hirsch is the W. J. Usery Chair of the American Workplace, Department of Economics, Andrew Young School of Policy Studies, Georgia State University, Atlanta, Georgia. He is also Research Fellow, IZA (Institute for the Study of Labor), Bonn, Germany. His e-mail address is (bhirsch@gsu.edu).
preferences of the average or median voter (that is, union member) rather than the mobile or marginal worker to whom employers are most responsive in competitive labor markets. Unions seek to increase and compress compensation, improve working conditions, and influence workplace jurisprudence. As compared to nonunion establishments, union governance is highly formalized, with conditions of work and pay designated in contracts determined through collective bargaining. Bargaining power is derived from legal rights granted in the National Labor Relations Act (NLRA) and from the economic environment in product and labor markets. Competition constrains union power by increasing the elasticity of labor demand (the employment-wage trade-off). Close substitutes in product markets limit the ability of firms to pass labor cost increases through to consumers. In factor markets, the ability to substitute capital or nonunion workers for union labor, say, through outsourcing of inputs or moving production to nonunion plants, likewise limits bargaining power.

Much of the economics literature on unionism adopts the “two faces of unions” approach originated by Freeman and Medoff. A comprehensive retrospective on the body of work sparked by their 1984 book—*What Do Unions Do?*—is offered in Bennett and Kaufman (2007). The “good” collective-voice/institutional-response face emphasizes union enhancement of voice, improvements in productivity, and decreases in earnings dispersion. The “bad” monopoly face emphasizes allocative inefficiency due to union wage distortions, along with decreases in productivity due to union work rules. Although theory is informative, the effect of unions on compensation, productivity, profitability, investment, and other dimensions of performance is largely an empirical question. This paper will present U.S. evidence pointing to substantial union wage premiums. Far less evidence is available on union effects on firm economic performance. My assessment of the U.S. evidence is that unions have, at most, small positive (but variable) effects on productivity insufficient to offset the substantial compensation gains, thus leading to lower firm profitability. Unionization is associated with lower investment in physical and intangible capital and slower growth (Addison and Hirsch, 1989; Hirsch, 2007b). Because union effects are the joint product of union and management interaction, workplace outcomes vary with the labor relations environment (Kleiner, Leonard, and Pilarski, 2002; Krueger and Mas, 2004).

Although corporate shareholders and unions have different views as to how a firm’s revenues are to be divided, they have a mutual interest in maintaining employers’ financial health. Thus, the sharp decline seen in private sector unionism might at first seem puzzling. To understand this shift from a union to nonunion governance norm, I begin with an overview of the rise and fall of unionism in the U.S. economy. Explanations for union density decline fall into three broad categories: structural, competitive, and institutional. I emphasize the competitive explanation. Unions in effect tax company returns via wage increases not offset by productivity gains. Union wage premiums are shown to be sizable, with only modest declines as union density has fallen. Union governance is formalized and deliberative. The combination of a union tax and sluggish governance is proving debilitating in economic environments that are highly competitive and dynamic.
Maintenance of wage premiums, coupled with sluggishness in the face of change, is illustrated below in two large, dynamic, and highly unionized U.S. industries: automotives and airlines. I conclude that collective bargaining is likely to remain a minority model, as nonunion norms of employee governance evolve in response to market forces and public policies. Union decline results in what is arguably an underproduction of worker voice and participation in the workplace. A concluding section peers briefly into what might or should emerge in place of traditional unions.

The Rapid Rise and Pervasive Decline of Private Sector Unionism

The Rise

Labor unions did not seem to have a bright future in the early 1930s. In his 1932 AEA presidential address, labor economist George Barnett (quoted in Kaufman, 2002, p. 330) stated: “It is hazardous to prophesy, but I see no reason to believe that American trade unionism will so revolutionize itself within a short period of time as to become in the next decade a more potent social influence . . . . We may take it as probable that trade unionism is likely to be a declining influence in determining conditions of labor.” However, a combination of forces led to a sharp increase in union density during the late 1930s and 1940s, peaking in the 1950s. As Figure 1 shows, private sector union density rose from about 12 percent in 1929 to 24 percent by 1940, and to 35 percent by 1947.

The rapid rise of U.S. unions suggests that sharp increases in union density are not likely to occur gradually, but in spurts as the result of major economic, social, and political upheaval (Freeman, 1998). Passage of the National Labor Relations Act (NLRA) in 1935 provided the legal and administrative framework that facilitated the rapid transition to an industrial U.S. economy in which union governance was the norm. A more fundamental explanation, however, was that the Great Depression was widely viewed as a failure of capitalism and the product of destructive competition. The result was a set of corporatist New Deal policies in which business, labor, and government were economic partners. Major industries—coal, steel, automotive—became unionized over a brief period, a transition encouraged by the Roosevelt administration in the 1930s and later reinforced by the industrial buildup for World War II.²

Following World War II, high inflation and strike activity shifted majority opinion toward support for greater limits on union power. Accordingly, the Taft-Hartley Act in 1947 outlawed union practices like closed shops and secondary boycotts, allowed states to pass “right-to-work” laws, and gave the federal govern-

² Wachter (2007) argues that the NLRA set up the administrative machinery for union organizing and governance, but constituted a break from the cooperative corporatist framework envisioned by the New Deal. The NLRA recognized collective bargaining as an adversarial system that would operate within a market economy. In this view, the NLRA planted the seeds for union decline, allowing union power to be constrained and eventually marginalized by competitive pressures.
Figure 1
U.S. Private Sector Union Density, 1929–2006

Sources: There is no definitive, fully time-consistent series on union density. Figures for 1929–1972 were compiled by Troy and Sheflin (1985) from union financial reports; those for 1973 forward are compiled from CPS household data (Hirsch and Macpherson, 2003, updated at (http://www.unionstats.com)).

Notes: I adjust 1973–76 CPS figures to account for association members, who are included both in the Troy-Sheflin series and in the CPS beginning in 1977.

ment the power to block or end strikes that might have national safety or health implications. Union density was largely flat through the mid-1950s, with peak density at 36 percent in 1953 and 1954. Private sector union density edged down slightly during the late 1950s and 1960s, and then began its long-term decline in the 1970s.

A Taxonomy of Decline

The decline of American unionism has been gradual but unrelenting. Private sector union density was 24.5 percent in 1973, 16.5 percent in 1983, only 11.1 percent by 1993, and 7.4 percent in 2006. The number of private sector union members was 15 million in 1973, a level roughly maintained through the end of that decade, but which had fallen to 8 million in 2006. As private union membership fell by nearly half, nonunion private wage and salary employment more than doubled from 47 million in 1973 to 108 million in 2006.

It is sometimes argued that the U.S. economy has deindustrialized, with manufacturing becoming an increasingly tiny sector. There is some truth in this statement. Manufacturing employment declined from 20.1 million in 1973 to 15.6 million in 2006, and the share of manufacturing to total private wage and salary employment has fallen even more sharply, from 32.5 percent in 1973 to 14.6 percent in 2006. However, nonunion manufacturing employment remained relatively steady over time, as union employment dropped sharply. As seen in Figure 2A, nonunion employment in manufacturing rose by about 1.5 million between 1973 and 2006, despite a 2.5 million decline after 2000. Union members employed in manufacturing fell from 7.8 million in 1973 to only 1.8 million in 2006.

Manufacturing is not typical of all sectors of the economy. After all, developed economies typically exhibit a declining employment share in goods-producing industries and an increase in service industries. But throughout much of the private
Figure 2
Union and Nonunion Employment in the U.S. Private Sector

Notes: Values for 1982 are interpolated. Communications industries shown in Figure 2C were expanded in 2003 to include publishing, libraries, and other information services.
nonmanufacturing economy as well, nonunion employment has displayed robust
growth, while union employment has been stagnant. As seen in Figures 2B and 2C,
such a pattern is evident in what historically were highly unionized construction
and transportation/communications/utility sectors. In construction, nonunion
employment rose from 2.5 to 7.3 million between 1973 and 2006, whereas union
employment fell from 1.6 to 1.1 million (corresponding to a decline in density from
39.5 to 13.0 percent). Between 1973 and 2006, private nonunion employment in
the transportation/communications/utility sector rose from 2.2 to 6.8 million,
while union employment fell from 2.3 to 1.6 million (corresponding to a density
decline from 51.4 to 19.0 percent).

A similar pattern is seen for all remaining sectors of the private economy,
shown in Figure 2D. At the same time that private nonunion employment grew
sharply, from 29.8 to 71.9 million between 1973 and 2006, union employment was
essentially flat, rising from 3.4 to 3.5 million (and density fell from 10.2 percent to
4.6 percent). Thus, the entire decline in private sector membership between 1973
and 2006 is accounted for by the three historically high union sectors shown in
Figures 2A-C, with most of the decline concentrated in manufacturing.

Although the focus of this paper is on the U.S. private sector, union members
are increasingly in the public sector. Trends in private and public union member-
ship and density between 1977 and 2006 are shown in Figure 3. As private sector
density declined from 21.7 to 7.4 percent, public sector density changed little, from
32.8 in 1977 to 36.2 percent in 2006 (15.9 percent of all wage and salary employees
in 2006 worked in the public sector). In 1977, only a quarter (25.8 percent) of U.S.
union members were public sector employees; by 2006 nearly half (48.0 percent)
were public employees. Collective bargaining in the public sector is not covered by
the NLRA, but by state and federal statutes largely enacted since the 1960s. In
comparison to private companies, the public sector is less subject to competitive
pressures.

Understanding the Decline in Private Sector Unions

There is no shortage of "suspects" to explain the decline in private sector
unionism, and no satisfactory means to distinguish each factor's contribution. I
group the suspected factors into three not fully distinct categories: structural,
competitive, and institutional. Structural explanations emphasize shifts in employ-
ment away from occupations, industries, and regions where union density has
traditionally been high toward sectors with lower density. The competitive expla-
nation, emphasized in this paper, focuses on the effect of unions on the financial
performance of companies and the competitive environment in which companies
operate, an environment affected by international trade, government regulation
(and deregulation) of product markets, and technological change. The institu-
tional category is a bit of a catchall, emphasizing factors that affect union organiz-
ing. I include the legal environment in which unions organize, the role of man-
Figure 3

Notes: Values for 1982 are interpolated. I begin with 1977 rather than 1975 to obtain a time-consistent definition of membership, which most affects public employees.

Management opposition, worker preferences toward unions, and government regulation of the labor market.

Structural Explanations

Historically, production workers in manufacturing located in north-central states were highly unionized. But over time, the shares of jobs in manufacturing industries, in production occupations, and in north-central states have declined. Such structural explanations for union decline lend themselves to measurement. To illustrate, I assess the structural explanation by classifying jobs by detailed industry and occupation and then asking the following question: had within-sector union density remained constant over time while the composition of jobs was allowed to change, by how much would union density have declined? This accounting exercise allows one to decompose union decline into that due to a change in job type and that due to lower density within job types.\(^3\)

I provide a decomposition of union density change between the years 1983 through 2002, a period where time-consistent codes for detailed industry and occupation can be readily constructed (Hirsch and Macpherson, 2007, tables 7b and 8b). After omitting industries and occupations in which there are no private sector workers and combining tiny groups, there are 211 industry and 366 occupation categories. Results are presented in Table 1. The first set of rows show that

\(^3\) Dickens and Leonard (1985) and Farber and Western (2002) provided earlier “accounting” analyses focusing on rates of union organizing and job creation/destruction.
### Table 1
A Decomposition of Union Density Change between 1983 and 2002

<table>
<thead>
<tr>
<th></th>
<th>Industry analysis</th>
<th>Occupation analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) Economywide union density</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1983</td>
<td>16.49</td>
<td>16.49</td>
</tr>
<tr>
<td>2002</td>
<td>8.60</td>
<td>8.60</td>
</tr>
<tr>
<td>Change in union density</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>7.89</td>
<td>7.89</td>
</tr>
<tr>
<td>2) Change in union density from 1983 to 2002 due to changes in union density within industries or occupations, holding employment constant at its 1983/2002 average</td>
<td>-6.29</td>
<td>-6.46</td>
</tr>
<tr>
<td>3) Change in union density from 1983 to 2002 due to changes in employment within industries or occupations, holding union density constant at its 1983/2002 average</td>
<td>-1.59</td>
<td>-1.43</td>
</tr>
<tr>
<td>4) What union density would have been in 2002 if union density changed within industries or occupations, but employment was constant</td>
<td>10.19</td>
<td>10.03</td>
</tr>
<tr>
<td>5) What union density would have been in 2002 if employment within industries or occupations changed, but union density was constant</td>
<td>14.90</td>
<td>15.06</td>
</tr>
<tr>
<td># of industry/occupation categories</td>
<td>211</td>
<td>$66</td>
</tr>
</tbody>
</table>


Private sector union density decreased from 16.5 percent in 1983 to 8.6 percent in 2002, a decrease of 7.9 percentage points. In the remaining rows, this change in density, $\Delta \frac{\%U}{i}$, is separated into that due to the change in density within each industry $i$ (or occupation $j$) and that due to changes in employment shares across industries (occupations), the latter interpreted as “structural” change.4

Results are clear-cut. As seen in the first column, line 5, if only employment across industries (that is, job structure) had changed between 1983 and 2002, with union density within industries remaining the same, then economywide union density would have fallen only 1.6 percentage points to 14.9 percent—just 20 percent of the total 7.9 percentage point decline in union density. If instead the structure of employment across industries had remained constant while union

---

4 Specifically, the decomposition is as follows:

Industries: $\Delta \frac{\%U}{i} = \sum_{i} (\text{Emp}_{i} \Delta \%U_{i}) + \sum_{i} (\%U_{i} \Delta \text{Emp}_{i}) = -6.3 - 1.6 = -7.9$

Occupations: $\Delta \frac{\%U}{i} = \sum_{i} (\text{Emp}_{i} \Delta \%U_{i}) + \sum_{i} (\%U_{i} \Delta \text{Emp}_{i}) = -6.5 - 1.4 = -7.9$

Emp$_{i,0}$ is the employment share in industry $i$ (occupation $j$); $\%U_{i,0}$ is union density in industry $i$ (occupation $j$). The first term on the right-hand side of the equation reflects the contribution of within-industry (occupation) changes in density, weighted by employment shares, to the total change in union density. The second term on the right-hand side of the equation reflects the contribution of structural employment changes, weighted by union density.
density within each industry changed as observed (line 4), then economywide union density would have fallen 6.3 percentage points to 10.2 percent—80 percent of the actual percentage point decline. Results are quite similar using occupation, as seen in column 2. Structural change in employment is not the principal reason for union decline.

Three minor caveats are added. First, the precise results of the decomposition are not unique. For example, results change modestly if a different employment or union density base is used (I use the average for the two years 1983 and 2002) or if one had a greater or lesser degree of industry and occupation aggregation. Second, some of the decline in density attributed to job structure might be attributed to the competitive and institutional explanations. For example, changes in technology, trade, regulation, immigration, and the like cause changes in industry, occupation, and job location. Third, previous changes in unionism from structural (or other) changes beget future changes. To some extent, unionization is an “experience good” (Gomez and Gunderson, 2004) whose spread is “contagious” (Holmes, 2006). Those localities seeing unionization of coal mines and steel mills in the past are now more likely to have unionized supermarkets and hospitals, as past unionization both creates an organizing infrastructure and increases workforce demand for representation. Conversely, declining levels of union density make future organizing more difficult.

**Competitive Explanations**

The competitive explanation is that union strength developed through the 1950s was gradually eroded by increasingly competitive and dynamic markets. To the extent that high union labor compensation is not offset by greater productivity or higher product prices, union gains can be thought of as a “tax” on firm profits. The competitiveness of the product market affects the ability of unions to acquire gains for their members. When much of an industry is unionized, firms may prosper with higher union costs as long as their competitors face similar costs. When union companies face low-cost competitors, labor cost increases cannot be passed through to consumers. Factors that increase the competitiveness of product markets—increased international trade, product market deregulation, and the entry of low-cost competitors—make it more difficult for union companies to prosper.

Readers probably need little convincing that the U.S. industrial sector operates in a highly competitive and dynamic environment. Economywide and manufacturing concentration ratios for value-added, employment, and payroll have remained steady or decreased over the last 50 years (White, 2002). Competitive pressures from international trade are strong: the value of imports as a percentage of GDP increased from 5.4 percent in 1970 to 17.2 percent in 2006:3 (U.S. Council of Economic Advisers, 2007, table B-1). Dynamism is evidenced by high rates of productivity growth and job churn. Output per work hour in the nonfarm business sector of the economy has doubled since 1970, from an index (with 1992 = 100) of 68.0 in 1970 to 137.7 in the third quarter of 2006 (table B-19). Productivity growth in manufacturing exceeds that economywide. In the U.S.
labor market between 1990 and 2005, the private sector (manufacturing sector) had a job destruction rate of 7.6 (5.3) percent per quarter and job creation rate of 7.9 (4.9) percent, implying a net employment growth rate of 0.3 (~0.4) percent (Davis, Faberman, and Haltiwanger, 2006, table 2A). To maintain union density as employment grows, unions must organize a number of existing and newly created jobs that exceeds the number of union jobs destroyed.

U.S. evidence on unions and firm performance bolsters the competitive hypothesis. Unions have a small average (and highly variable) effect on productivity, failing to offset higher compensation. Union coverage thus leads to lower profits, based on accounting measures such as earnings/sales or earnings/asset ratios or on market-based measures (market value divided by capital costs or sales). For example, in Hirsch (1991b) I find that compared to nonunion manufacturing companies, profitability is lower by about 5 percent for otherwise similar companies with low union coverage, 15 percent for those with medium coverage, and 20 percent for those with high coverage. There is no reason for stock market returns to be systematically lower for union than for nonunion companies, since expectations of lower earnings decrease a firm’s share price until risk-adjusted expected returns equalize across firms (for evidence, see Hirsch and Morgan, 1994).

Union companies in the United States also tend to have lower investment in physical capital and in innovative capital (R&D), higher debt, and lower growth (Hirsch, 1991a, 2007b). The capital investment result is particularly interesting. All else the same, high wages increase incentives to invest in labor-saving capital. However, lower profits reduce internal investment funds and union wage premiums effectively tax the returns accruing to investment in immobile, long-lived capital (Grout, 1984; Hirsch, 1991a; Hirsch and Prasad, 1995; Addison and Chilton, 1998). Absent the ability of a union to promise in a credible manner that future returns will not be taxed, a union company may decrease investment.

Evidence directly linking union decline to competition is limited, although Slaughter (2007) shows that union density is lower in manufacturing industries with a high degree of global engagement, in particular those with inward foreign direct investment. Slaughter’s interpretation is that increased capital mobility has raised labor demand elasticities and weakened union bargaining power. An important

---

5 Empirical work on performance is severely limited by the absence of publicly available information on the union coverage of U.S. firms. Our knowledge relies heavily on studies that collect union data from firms and link it to publicly available data (Hirsch, 1991a) or to analyses using nonpublic surveys (Clark, 1984). Although U.S. studies probably do a good job uncovering cross-section differences in firm outcomes with respect to union coverage, holding constant other measurable determinants, it is not clear how well these studies identify causal effects. (An exception is DiNardo and Lee (2004), who show that establishments in which workers barely approved union coverage have outcomes little different from establishments in which coverage was barely rejected. It is not clear whether these results for recent, barely-approved unions can be generalized to the larger population of union establishments.) The fact that union companies often perform poorly is itself informative, even if we do not understand fully the causal mechanisms at work.

6 In Hirsch (2007b), I survey U.S. evidence on unions and firm performance. Fuchs, Krueger, and Poterba (1998) examine the views of labor economists at top universities. They asked the following question: “What is your best estimate of the percentage impact of unions on the productivity of unionized companies?” The median response was zero and mean response 3.1 percent.
Implication of the competitive thesis is that a highly competitive economy not only constrains union density, but also limits the economywide costs of unionism since resources are more readily reallocated to sectors with the highest expected returns.

**Institutional Explanations**

The “institutional” category is intended to capture factors that affect the conduct and outcomes of union organizing. Included in this category are worker sentiment for union representation, management opposition to unions, the interpretation and enforcement of labor law, and governmental workplace regulations that may substitute for or complement union representation.

A frequent argument is that unrepresented workers have substantial unmet demand for representation, but fierce management opposition prevents this from being translated into union coverage. Survey evidence indicates strong support for union organizing among noncovered workers (Freeman and Rogers, 1999). There is a large management consultant/union avoidance industry that helps companies avoid the organizing of unions (Logan, 2006). Employers are frequently accused of unfair labor practices, the most frequent charges being refusal to bargain once workers vote for representation, and illegal discharge or other discrimination against employees for union activity. Expected penalties for engaging in unfair labor practices are low. At the workplace, supervisors have greater access to workers than do unions, although the Internet has lowered unions’ cost of communicating with workers. Typically, the longer an organizing campaign runs prior to voting, the lower is union support.

Yet there is reason for skepticism that the unmet demand for union coverage is large. Surveys show that workers want not only greater participation in workplace decision making, but also a cooperative rather than adversarial environment (Freeman and Rogers, 1999). The demand for increased participation is similar for union and nonunion workers, suggesting not only that workers want voice, but also that many unions are not providing the voice that workers want.

Currently workers vote for union representation through use of a secret ballot in NLRB-supervised elections. Though proposed reforms would mandate union recognition following a “card check”—signatures from a majority of workers supporting the union—I am reluctant to believe that card checks or answers to surveys about union support more accurately reflect worker preferences than do secret ballot results.\(^7\) Still, demand for some kind of voice is not being met through the current system. Bryson and Freeman (2006) conclude that underlying worker preferences are similar in the United States and United Kingdom, but that workplace outcomes differ because the United Kingdom provides a greater range of institutional options.

---

\(^7\) The Employee Free Choice Act of 2007, passed by the House in March 2007 but killed in the Senate in June 2007, would have required an employer to recognize a union once a majority of workers sign cards authorizing union representation. Under current law, employers may choose to recognize a union based on such a card check or demand a secret ballot election. Passage of such an act, presumably by a Democratic Congress and president, would increase union organizing.
Managerial opposition to organizing has been a key factor reinforcing union decline. Management’s dislike of unions is partly ideological, but if unions reduce employer discretion and company profits, as the evidence suggests, resistance is to be expected. Increased opposition may result in part from increasingly competitive markets. If labor laws and their interpretation by the NLRB were more favorable to organizing, union density would be somewhat higher. Yet density has declined for decades, with NLRB majorities that are Democratic as well as Republican. Public policy should promote enhanced workplace voice and cooperation that adds value (defined broadly) to society. It’s not obvious that unionism in its current form is the preferred path toward this goal.

A final environmental consideration is the role of government mandates and labor regulations affecting most workplaces. Government protections from discrimination, safety hazards, pension failure, and the like may lessen worker demand for unionism. Neumann and Rissman (1984) provide time-series and cross-section evidence consistent with this substitution thesis.

Of the three explanations for union decline, the competitive thesis is the most fundamental. A large union sector requires that existing union establishments prosper and that large numbers of new jobs “born” nonunion be organized. The more competitive are product and financial markets, the more critical it is that unions not have deleterious effects on firm performance. Competitiveness is a driving force behind the structural and environmental categories as well. Competition and new technologies help determine the structure of jobs (their type and location). Competitive pressures influence union organizing through effects on management opposition, worker sentiment, and public policies.

**Union Wage Premiums in the U.S. Private Sector**

Union pay is determined by a collective bargaining process shaped by the preferences of union members and the bargaining power of the parties, the latter influenced by product and labor market conditions. Union leaders are elected and union contracts must be approved by a majority of rank-and-file. Basic models of union behavior treat the decisions of union leaders as responsive to the preferences of the median voter or member (Farber, 1986, provides a survey). Members face a trade-off between wages and employment (although settlements need not be on the labor demand curve), with the trade-off influenced by the ability of firms and customers to substitute to and away from union workers, union establishments, and union-made goods. Incumbent workers may place low weight on employment opportunities that could exist at a lower wage, while median (often older) members may feel insulated from layoffs based on seniority, except when establishment closings or firm bankruptcies are a threat.

---

8 A union practitioner (Sleight, 2007, p. 603) states: “[U]nions need to have the assent of their members to make significant changes. The very problem with unions that F&M [Freeman and Medoff] identify, monopoly wages, is exactly what our members want!”
In contrast to the limited information available on company unionization and economic performance, data on wages are readily available. These data permit estimation of union wage premiums—the percentage difference in the wages of similarly skilled union and nonunion workers in similar jobs. Lewis (1986) distinguishes between union wage gaps and wage gains. A gap is equivalent to a premium, as defined above. Wage gains measure union wages relative to wages that would exist absent unionization, thus taking into account union effects on nonunion wages. The empirical literature suggests the net effect of unions on nonunion wages is small, but positive. (If zero, wage gaps and gains would be equivalent.) Changes in union wage gaps (premiums) over time provide a rough but useful measure of changes in the costs of unionism to firms. Although the focus here is on wages, it is total compensation (wages plus benefits) that is the more relevant measure. Evidence on benefits is limited, but data available points to a union benefits premium that exceeds the wage premium (Budd, 2007).

In Hirsch and Macpherson (2007, table 2), my coauthor and I provide time-consistent regression estimates of union–nonunion wage premiums (gaps) for the years 1973–2006. We use the May Current Population Survey (CPS) earnings files for the years 1973–81 and the monthly CPS earnings files for 1983–2006. Estimates of the union gaps are obtained from annual wage equations, where the natural log of the wage is the dependent variable, while union status and controls for worker/job characteristics are included as independent variables. The regression coefficient on union status is the measure of the union wage premium, as shown in Figure 4. Because wages are measured in natural logs, the coefficient on union status can be interpreted as a proportional difference. Controls include years of schooling, potential years of experience and its square (interacted with gender), and variables for marital status, race and ethnicity, gender, part-time work, large metropolitan area, public sector, region, broad occupation, and broad industry (omitted for industry-specific estimates).

Before discussing results, it is important to mention two measurement issues. The estimates shown in Figure 4 assume that union status is exogenous; that is, determined independently of the wage. It does not account for selection into union status, a two-sided process determined by choices of individuals to be in a union job queue and choices of union employers to hire applicants from the queue (Abowd and Farber, 1982). This selection process leads to positive selection for skills among workers with low measured credentials and negative selection for skills among those with high measured credentials. For example, high school dropouts who are union members tend to be more able than average dropouts (because union wages are high enough that the firm need not hire workers with low skills), while union members with a college degree tend to be less able than average college graduates (because most college graduates, because of their

---

9 The log differential is an approximate proportional difference, the union minus nonunion wage \((W_u - W_n)\) divided by some “average” wage between \(W_u\) and \(W_n\). The percentage differential \([\{(W_u - W_n)/W_n\}]\) is typically approximated by \([\exp(d) - 1]100\), where \(d\) is the log differential.
Figure 4
Private Section Union Wage Premiums, 1973–2006


skills, can command relatively high wages in the nonunion sector). Although standard union premium estimates for dropouts will be overstated and for college graduates understated, the average union wage gaps shown in Figure 4 are not greatly affected by selection (Card, 1996; Hirsch and Schumacher, 1998; Hirsch, 2004).

A second issue concerns treatment of “imputed” earnings (Hirsch and Schumacher, 2004; Bollinger and Hirsch, 2006). The CPS monthly earnings files since 1979 include “allocated” earnings in which the Census imputes values for those who refuse or are unable to report earnings, once about 15 percent of earnings records and now about 30 percent. Imputed values are obtained by assigning to each worker with missing earnings, the earnings of a matched “donor” with an identical set of selected characteristics (gender, plus broad categories for age, race, schooling, occupation, and hours worked), but not union status. In Hirsch and Schumacher (2004), my coauthor and I show that inclusion of imputed earners creates a “match bias” for characteristics like union status—an attenuation of CPS wage gap estimates with respect to non-match characteristics of about 25 percent in recent years. This bias exists for union wage premium estimates and for wage gap estimates with respect to characteristics like industry, foreign-born, citizenship status, Hispanic, marital status, city size, and many other attributes. The Hirsch–Macpherson estimates shown in Figure 4 exclude imputed earners.

Have private sector union wage premiums declined over time as density and union
organizing strength have diminished. The short answer is yes, but not by much. Readily evident in Figure 4 (see the line with the squares) is a large but fairly stable private sector union wage premium for over 20 years, but one displaying a modest downward trend from a 1984 peak of .249 to a 2006 gap of .189 (public sector gaps are smaller). Closer inspection suggests the premium is (with a lag) countercyclical (Blanchflower and Bryson, 2003), not surprising given union wage setting with multiyear contracts. An unresolved puzzle is that Bureau of Labor Statistics establishment data produce relative union–nonunion rates of wage growth not fully consistent with CPS household data (Hirsch, Macpherson, and Schumacher, 2004).

Figure 4 also provides union premium estimates for three traditionally union sectors: manufacturing, transportation/communications/utilities, and construction. Union density in each of these sectors has fallen sharply, but union members in these sectors continue to realize substantial wage premiums. By the mid-1990s, the premium in manufacturing nearly caught up with the total private sector average. Union workers in transportation, communications, and utilities realize a more substantial premium, one largely maintained despite substantial deregulation and increased competition. Union–nonunion wage gaps in construction are far larger than for the private sector as a whole, but display a downward trend. Some portion of the union gap in construction is probably not a true premium, instead reflecting unmeasured union–nonunion differences in occupational skill not directly measured in the CPS.

In short, union members continue to receive a solid wage advantage compared to similar nonunion workers, despite the declining role of unions in the U.S. private sector. The union premium is not very different today from that seen in the 1970s and is substantially larger than in most developed economies (Blanchflower and Bryson, 2003, and included references). Because the typical union workplace does not generate value-added sufficient to offset higher union compensation, the union premium reinforces the long-run decline in private sector unionism.

Unions Out of Time: A Sluggish Institution Confronts Dynamic Markets

In a less dynamic environment, a sluggish union governance process could provide benefits like worker democracy and collective voice with low costs. But in a dynamic environment, union governance is less likely to survive the market test. Increasingly, the workplace norm is one of nonunion governance with constrained employer discretion and the use of modern-day human resource management practices.

---

10 There is no standard measure of union bargaining power. Were firm-level (inverse) labor demand elasticities available, they would be used. Because of data availability, union density and wage premiums are often used as measures. Cramton and Tracy (1998) construct a model in which weak bargaining power leads to less use of strikes and greater use of holdouts (working with an expired contract). Bargaining power fell sharply during the 1980s, as measured by lower use of strikes relative to holdouts. They conclude that much of the decline was due to employers’ increased willingness to use replacement workers.
Unions formalize workplace governance, rely on explicit contracts, and limit managerial discretion. Although collective bargaining contracts permit considerable employer discretion in the daily operation of a workplace, substantive changes in wages, benefits, or working conditions must be negotiated with the union. Typically these changes are negotiated when, every three years or so, an expiring contract is to be replaced. Formal contracts have advantages, among them limiting opportunistic behavior by employers. But they come with a cost, reducing the ability of employers to make needed adjustments in the face of unanticipated shocks.

Wachter (2004) provides an insightful analysis of the labor contracting relationship, comparing the use of union governance with employer discretion. All workplaces must deal with four fundamental factors affecting labor contracting: match-specific investments, risk bearing, asymmetric information, and transaction costs. As described by Wachter, both union governance and constrained managerial discretion can solve problems dealing with the first three factors, but union firms are disadvantaged by high transaction costs. I argue here that in dynamic economic settings, the high transaction costs of union governance result in sluggish response to shocks, placing union companies at a competitive disadvantage (also see Hirsch and Hirsch, forthcoming).

Costly transacting in union workplaces stems in part from the adversarial nature of traditional unionism and the limited opportunity for worker voice outside of formal union channels. Embedded in the National Labor Relations Act is a view that workplaces rely on top-down control moving from managers and supervisors (who cannot be unionized under the NLRA) to workers with jobs requiring minimal decision making and discretion. Job hierarchies, however, are not as clear-cut as envisioned in the NLRA. Technology is increasingly eliminating production and service jobs with “routinizable” or programmable tasks (Autor, Levy, and Murnane, 2003). Worker decision making, cooperation, and identification with one’s employer, rather than with an adversarial agent, are essential elements of performance in most workplaces. Research in human resource management and personnel economics (Lazear, 1995) describe the panoply of pay and evaluation systems structured to align the incentives of workers with those of the firm. In contrast, union workplaces continue to rely on seniority-based pay and wage compression within and across positions. Verma (2007) surveys union effects on workplace governance.

Of course, union governance need not be this way. Collective bargaining agreements can include broad management rights clauses that provide employers with considerable discretion, as the National Labor Relations Board has held (St. George Warehouse, Inc., 341 N.L.R.B. 904 [2004]). Unions and management could adopt many features of modern pay and evaluation systems. More cooperative and less adversarial relationships between management and labor could arise. However, such relationships are the exception rather than the rule and appear destined to remain so due to the union political process and expectations of rank and file, coupled with union animus among many employers. The bottom line is that
existing governance of the employment relationship through collective bargaining is not surviving the market test.\footnote{The argument in this paper is that traditional unionism is a poor fit in a dynamic, competitive economy. In the macroeconomics literature, Blanchard and Giavazzi (2003), among others, provide related arguments in their attempt to understand the evolution of European economies. They emphasize the importance of competitive product markets, which limit labor market rents.}

**Case Studies: Motor Vehicles and Airlines**

Two industries where unionization has played a major role are the automotive industry (both motor vehicles and automotive equipment manufacturing) and the airline industry. Both industries have grown over the past 30 years, in contrast to many U.S. manufacturing industries. What union companies in these industries have in common is that in the face of economic shocks coupled with high contractual labor costs (including "legacy" costs from promised pension benefits and retiree health costs), needed adjustment to a cost-competitive economic environment has been sluggish, occurring only in the face of near or actual bankruptcy. Although bankruptcy provides a mechanism to ratchet labor costs closer to competitive levels, the process involves substantial costs and is often accompanied by an acrimonious labor relations environment unlikely to facilitate future growth. The two industries also have a key difference. In the automotive industry, union employment has declined sharply. By contrast, the airline industry remains highly unionized, and its unions retain unparalleled strike threat power.

**Motor Vehicles and Motor Vehicle Equipment Industries**

In many manufacturing industries, production has moved overseas to developing economies, particularly in industries where shipping costs are low relative to value, where plants do not require massive and sophisticated capital investments, and where work is low-skill intensive. Automobile assembly and equipment manufacturing, however, are concentrated in developed economies outside of central cities, with plants geographically clustered within a day's drive, thus lessening transportation costs, facilitating competition among suppliers, and enabling technology and human capital spillovers (Klier, 2005). The incentive for parts manufacturers to locate near vehicle plants has increased over time as just-in-time inventory and production methods have increased in importance. All else the same, it is beneficial to locate plants and production networks not too far from consumers.

Total U.S. employment in motor vehicle and equipment manufacturing has grown modestly over the past 30 years. But the overall growth masks dramatic employment shifts from U.S.-owned union establishments located in north-central states toward a mix of U.S. and foreign owned (primarily Japanese and European) establishments in locations further south with primarily nonunion workforces.

Shifts in union density can be seen using Current Population Survey data for 1973–2006, shown in Figure 5. Employment in the combined motor vehicle and
automotive equipment industries changed from 1.2 million in 1973 to 1.0 million in 1983, 1.1 million in 1990, 1.3 million in 2000, and 1.4 million in 2006. At the same time, union membership (and density) dropped sharply from 830,000 (71.0 percent) in 1973, to 590,000 (58.8 percent) in 1983, 540,000 (48.4 percent) in 1990, 470,000 (35.9 percent) in 2000, and just 360,000 (26.0 percent) in 2006.

Calculations from the Current Population Survey also reveal geographic shifts in employment. In the late 1970s, motor vehicle and parts manufacturing was concentrated in Michigan (41 percent of all workers in the industry), Ohio (12 percent), and several other high-union states. By 2004–2006, most of these states realized large employment share losses—Michigan (13 percentage points), New York (3.5), Wisconsin (1.5), and Ohio (1.4). Employment share gains were evident among mostly nonunion facilities in southern states like Tennessee, South Carolina, Alabama, Kentucky, all of which gained about 2–3 percentage points of the total automotive workforce between 1977–79 and 2004–2006. The Midwest exception was Indiana, increasing its automotive employment share from 6 to 8 percent, while union density declined from 70 to 28 percent, roughly mirroring the nationwide change.

Technological change in the automotive industry has been rapid, with substantial increases in productivity and decreases in quality-adjusted prices. Although U.S. companies fell behind their Japanese and European competitors, the productivity and quality gap has been sharply narrowed (Baily et al., 2005). But change has not been fast enough. In order for the Big Three to have prospered, they would have needed to maintain their market shares and avoid substantial price discounting; then they would have been able to spread their legacy retiree pension and health commitments over a larger number of automobiles and employees. This scenario did not play out. Instead, the Big Three U.S. companies lost market share (in particular, General Motors); they retained a commitment to pay retiree health
and pension costs; and their ability to trim labor costs remained limited because workers displaced due to technology or restructuring were paid while in a “jobs bank.” Vehicle manufacturers trimmed costs of inputs by putting substantial pressure on their auto parts suppliers to lower prices, at the same time demanding improved product quality and just-in-time delivery.

In the end, something had to give. Auto parts suppliers Delphi (a corporate spin-off of GM), Dana, Collins & Aikman, and others filed for bankruptcy, a process (in progress as this is written) through which worker compensation will be substantially lowered, plants will be closed, and some portion of Delphi’s legacy costs will be shifted to GM or repudiated. To reduce employment and shut down plants no longer needed, Delphi, GM, Ford, and others have structured various lump-sum buyouts of workers, the take-up rates on these offerings being sizable. Daimler-Chrysler has proposed spinning off its Chrysler automotive group to a private equity company with a reputation for cutting costs. The UAW (United Auto Workers) began contract talks with the Big Three in July 2007, a principal issue being how to cut company costs for what are currently generous employee and retiree health care benefits.

What might have happened had the Big Three not been unionized? Rather than “hitting the wall,” with bankruptcy (or its threat) producing the substantial employment and compensation decreases, a gradual adjustment process might have occurred. Over time, plants would have been closed, staffing levels decreased, and compensation growth slowed. As seen elsewhere in the private sector, employee health plans would have included considerable cost sharing and promises for retiree pension and health benefits would have been more constrained. Of course, there is no assurance that managers at the Big Three, absent the constraints of the union governance process, would have steered their companies toward healthy financial outcomes. What is clear is that the collective bargaining process cannot easily shield employees from long-run market forces or company-specific strategic failures.

The automotive industry is typified by large and medium-sized companies and plants. Whether union or nonunion (or U.S. versus foreign-owned), employment governance must be highly formalized. I am suggesting that adversarial union governance has proven a disadvantage relative to nonunion human resource management and greater managerial discretion, in particular when facing technological advances, domestic and international competition, market demand shocks, and the like. The tendency of union governance to be sluggish is not an economic law. Some union workplaces have well-functioning employee governance, good labor relations, and respond well to economic change. No doubt more workplaces would fit this characterization were not antipathy toward unions so widespread among U.S. managers. But adversarial union governance, both economywide and in an automotive industry characterized by a generally stable labor relations system, is failing the market test. To state that companies with union governance tend to perform poorly is not intended as criticism (or endorsement) of behavior among union leaders or management, but as a critique of the current U.S. labor relations system.
Airline Industry

The airline industry provides another example of union governance’s slow response to economic shocks. There is no large private sector industry where unions have higher density and greater bargaining power than in airlines. Roughly half of all workers in air transportation were unionized before deregulation in the mid-1970s and about half remain unionized today (Hirsch, 2007a). All major carriers, including Southwest, are unionized (Delta pilots but not other workers are unionized). Mid-size national and regional carriers include a mix of union and nonunion companies. Union density in the industry remains high, partly as a carryover from the regulatory period but, more fundamentally, because strong bargaining power makes representation attractive to workers. A profitable nonunion carrier paying well below other carriers would quickly be organized. Bargaining power is substantial because of the strike threat. A strike by a carrier’s pilots, flight attendants, or mechanics (and possibly other workers) can shut down all flight operations. A shutdown in a service industry can be particularly costly. Unlike consumer durables, transport services cannot be stored or shifted in time. Many customers can switch to non-struck carriers. Because shutdowns are so costly, strikes are rare, but unions are able to capture rents for their members.

What has emerged in the airline industry is compensation that reflects a “union tax” cycle. Union airline workers, particularly pilots, realize substantial wage premiums (Hirsch and Macpherson, 2000; Hirsch, 2007a). Following periods in which airlines have been relatively profitable, such as the late 1990s, union contracts “tax” those profits and wage premiums rise. Following substantial losses, unions provide contract concessions, as seen for some carriers in the mid-1990s and again during the last several years. But union response to changing economic conditions takes time. In the perfect storm that hit the airline industry in the early 2000s, the response was too slow. Adverse conditions faced by airlines included a recession hitting in 2001 as high contract wages were taking force; the 9/11 attacks and a 20 percent reduction in flights; a stock market downturn destroying pension wealth; Internet pricing that lowered carrier margins; increasing market shares of “low-cost carriers”; and later, increasing fuel prices. US Airways and United entered bankruptcy protection in 2002 (and US Airways again in 2004), while Delta and Northwest entered bankruptcy on the same day in 2005. American Airlines, the only legacy carrier that has not disappeared, been acquired, or entered bankruptcy since deregulation in 1978, was facing a bankruptcy filing in 2003 until it received concessions from its unions.

The case of Delta is instructive. Prior to 2001, Delta had a strong balance sheet with low indebtedness and, with the exception of its pilots, a nonunion workforce, although pay for nonunion workers had to be similar to that for union workers to deter union organizing. By 2003, Delta faced deteriorating product market conditions coupled with industry-leading pay for its pilots and other workers. Financial viability required that Delta sharply reduce labor and other costs, requiring substantial concessions from pilots. Despite its initial advantages, the company was unable to steer a path to financial viability, instead accumulating massive debt and finally resorting to use of the costly bankruptcy process to achieve lower costs.
Despite a long-run relationship between the major carriers and their unions, labor relations throughout much of the industry have remained highly contentious, with the exception of Southwest and, more recently, Continental. In most industries, companies that have high costs and respond slowly to economic shocks are likely to wither as customers switch to goods produced by domestic or foreign competitors. Emerging successfully from bankruptcy may not be a viable option for such companies. In the airline industry, however, bankrupt carriers have continued operations, retained much of their customer bases, and reemerged with revised business plans and lower labor costs and debt. Although price competition can limit the size of union wage premiums, union coverage in the industry has not been seriously threatened.

What Might Follow Private Sector Unions?

As long as product markets remain competitive and dynamic, American-style unionism will remain a minority workplace governance structure. For it to be otherwise would require both increased worker and voter desire for unionization, and changes in union governance that create value-added in the workplace. What we are likely to see in the near future is more of the same: an evolving private sector governance norm of management discretion, constrained by market forces, public policies, and societal norms. That outcome is hardly a terrible one. However, labor market norms and institutions surviving the market test need not be optimal. Worker voice may be underproduced in a competitive labor market (Levine and Tyson, 1990; Freeman and Lazear, 1995). Moreover, the governance structures that emerge and survive in the marketplace are determined in part by public policies with respect to the workplace, trade, immigration, pensions, health, taxes, and the like.

Policies that encourage the development of value-added worker voice and cooperation in the 90-plus percent of the private sector not unionized warrant debate (Estreicher, 1994; Levine, 1997; Freeman, 2005; Freeman, 2006; Hirsch, 2007b; Hirsch and Hirsch, forthcoming). As one example, modest changes in the NLRA would reduce current legal barriers to worker participation in nonunion workplaces. More radical proposals exist, although such labor law reform is barely visible on the political radar screen. “Conditional deregulation” would provide the option for companies with independent worker representation to waive selected workplace regulations. A far-reaching reform would change the current workplace governance default from one of nonunion to some new form including employee participation, a default from which employers and employee groups can move if there is mutual agreement to do so. Freeman (2005, 2006) would like to see labor law innovations at the state level, although under current court holdings, state regulations overlapping with the NLRA are preempted. The policies described

---

12 Sections 8(a)(2) and 2(5) of the NLRA prohibit employers from creating or supporting any worker group that engages in discussion with employers about grievances, disputes, wages, rates of pay, hours of employment, or conditions of work.
above might complement or provide substitutes for traditional unionism and governmental regulation of the labor market.

There currently exist several nonunion- and union-sponsored institutions that play a limited role in the labor market (Freeman, Hersch, and Mishel, 2005). For example, human rights groups, the living wage movement, and public interest legal organizations lobby on labor-related issues. There exist member-based groups such as professional organizations, groups organized around policy interests, and nonprofit community- or industry-based groups often focused on worker training. Freeman and Rogers (2002) discuss an “open-source” unionism wherein unions provide information and low-cost services to noncovered workers. More broadly, the Internet lowers communications costs for unions and other labor-related organizations (Freeman, 2005). I am not suggesting that any one of the above institutions will emerge as a dominant or even important player. The decline of traditional unions, however, may leave a partial vacuum in workplace governance, opening possibilities for new workplace institutions. In order for such institutions to take hold, they must flourish in the U.S. economic environment of open, competitive, and dynamic markets.

The author thanks Alex Bryson, Bruce Kaufman, Michael Wachter, and the editors of this journal for helpful comments and discussion.

References


