

University of Pennsylvania Carey Law School

Penn Carey Law: Legal Scholarship Repository

Faculty Scholarship at Penn Carey Law

5-12-2003

Judging Unions' Future Using a Historical Perspective: The Public Policy Choice Between Competition and Unionization

Michael L. Wachter

University of Pennsylvania Carey Law School

Follow this and additional works at: https://scholarship.law.upenn.edu/faculty_scholarship



Part of the [Economic Policy Commons](#), [Labor and Employment Law Commons](#), and the [Labor Economics Commons](#)

Repository Citation

Wachter, Michael L., "Judging Unions' Future Using a Historical Perspective: The Public Policy Choice Between Competition and Unionization" (2003). *Faculty Scholarship at Penn Carey Law*. 26.
https://scholarship.law.upenn.edu/faculty_scholarship/26

This Article is brought to you for free and open access by Penn Carey Law: Legal Scholarship Repository. It has been accepted for inclusion in Faculty Scholarship at Penn Carey Law by an authorized administrator of Penn Carey Law: Legal Scholarship Repository. For more information, please contact PennlawIR@law.upenn.edu.

Judging Unions' Future Using a Historical Perspective: The Public Policy Choice Between Competition and Unionization

Michael L. Wachter^{*}
University of Pennsylvania

The prospects for private sector unionism, in the context of the prolonged decline in membership since the early 1950s, have long captivated labor economists, and deservedly so. Today, private sector union membership is less than 10 percent of private sector employment, far below its peak of 36 percent in the early 1950s.¹ In the construction industry, a stronghold of private sector unions, union employment share is down from over 80 percent to approximately 20 percent.² This unrelenting decline represents one of the most important institutional changes affecting the United States economy.

In this paper I address the future of private sector unions in a broad historical and institutional context. In doing so I make extensive use of a recently published collection of papers on the topic, "The Future of Private Sector Unionism in the United States," edited by James Bennett and Bruce Kaufman, which also adopted a broad focus in exploring the future of unions.³

The positions I develop are the following. First, although numerous issues have been cited for the decline in unions over the past three decades, in the end they all come down to a single question: the ability of unions to achieve their traditional goals in the economic environment they confront. I consider the primary goal of unions to be taking wages and working conditions out of competition, although in so defining the goal I am forcing a complex set of goals into a single statement.⁴ This, in turn, amounts to taking advantage of opportunities for obtaining economic rents or removing exploitative or monopsonistic elements from firms. Consequently, it is the continuing trend toward a more competitive United States economy that represents the critical economic factor in the multi-decade decline in unions.

Second, to predict the future of unions, understanding their rapid rise to prominence between the 1930s and the early 1950s is at least as important as explaining their multi-decade decline. More specifically, to predict the near-term future the recent past is sufficient, but to understand what factors might be needed for a major resurgence in unionization, it is necessary to go back to the earlier period of rapid union growth. What factors were in place at that time? I argue that the high-water mark of unionization, rather than reflecting inevitable economy-wide forces, was itself highly dependent on the convergence of historical events that promoted unions. These include the anti-competition policies adopted during the Great Depression and the prevailing view that unions improved the efficiency and functioning of labor markets, making them less exploitative. By the time union employment reached its peak, however, the economy, including the labor market, was becoming more competitive, and the adoption of human resource management policies by nonunion firms was seen as successful in removing exploitative elements from the workplace of individual firms.

This view of unions is supported by the nature of the National Labor Relations Act (NLRA) and particularly by the collective bargaining apparatus that it created and supports. The collective bargaining apparatus, although ideally suited for protecting rents or attacking monopsonistic forces, is not well suited for a highly cooperative employment relationship. Contracts are a high cost method for organizing the employment relationship inside firms. Where the firm and its employees can cooperate using norms--or non-legally enforceable rules and standards--they can operate more efficiently than can a firm where the employment relationship uses the more costly and less flexible contract mechanism.

The paper is organized along the following lines. In the first section I evaluate the evidence for the causes of the decline in unionization rate. The evidence appears overwhelming that if the institutional setting does not change, unions will remain less than 10 percent of the private sector workforce. In Sections 2 and 3, I discuss the longer run historical record and develop the point that the high-water mark of unionization fits the convergence of specific historical events. My discussion relies extensively on the accounts being written as the events were taking place. More important than a modern reinterpretation of the past, we need to know what the actors, at the time, actually thought was happening. In Section 4, I examine the improvements in human resource management and what those improvements imply for the union sector. In Section 5, I examine the costliness of the NLRA collective bargaining apparatus. The last section summarizes the paper and draws some conclusions.

I. Economic Factors in the Decline in Private Sector Union Employment

There is a virtually unanimous view that assuming no major changes in the institutional environment, union density in the private sector will remain below 10 percent, and perhaps fall to close to 5 percent. The major reasons for this have been extensively documented in the literature.⁵

First and foremost, the American economy has become much more competitive over the past 50 years. Factors include the growth of the internationally traded goods sector; the increase in multinationals with multi-country production loci such as European and Japanese car companies with plants in the United States; the efficiency of capital markets in providing seed capital for new ventures to enter industries that earn super-competitive returns; and the decline in the number of industries that are regulated where government rules have protected firms from competition. The result is that few sectors of the economy are insulated from competition. As a consequence, these sectors cannot serve as a source of reliable rents for business and labor to negotiate over.⁶

Increased competition produces more winners than losers, but it does produce losers. Among the losers are the companies and their employees with cost structures that are above those of competitors and potential entrants. Losses are also suffered by companies and their employees who are being deregulated and who survived in the past

because of government protection from competition. To the extent that a primary goal of unions is to take wages and working conditions out of competition, they become primary losers as well.

Second, the overwhelming evidence supports the contention that unions do indeed raise wages above competitive levels. For a time, some researchers contended that the increase in union wages reflects an increase in productivity, but there has never been any reliable evidence to support this conjecture. The result of the wage premium is to put unionized firms at a competitive cost disadvantage. All else equal, one would expect that the union compensation premium would decline as industries became more competitive. In fact, the reverse occurred, particularly over the 1970s and early 1980s. Although the union wage premium has declined from historically high levels reached in the early 1980s, the premium or percentage wage differential is still above levels existing in earlier decades. In other words, while the American economy is much more competitive than say in the 1950s or 1960s, union wage premiums are higher today.⁷

Third, the private-sector American economy is becoming increasingly an economy where job creation occurs in industries and occupations and among nonunion firms and among demographic groups where unionization rates have historically been low. Based on the accounting methodology of Farber and Western (2002), American jobs are disproportionately being created in the service-producing sectors rather than the goods-producing sectors, and in professional and managerial occupations rather than in blue-collar occupations. In addition, the American labor force is increasingly better educated and more mobile, whereas unions have traditionally been strongest among less-educated and less-mobile workers.

The compositional effects are partly exogenous, but each of the elements leverages the effects of competition and the rising union wage premium. Although the growth in service sector employment is largely an exogenous factor, most service sector-producing industries are highly competitive. As the economy becomes more service oriented, it becomes more competitive at the same time.

But the high wage premium also contributes to the compositional shift between union and nonunion firms within a given sector. At least partly as a result of the higher costs, R&D expenditures and new capital investments in unionized firms lag behind levels in nonunion firms. Consequently, new employment opportunities, whether through the construction of new establishments or the startup of new businesses, occurs primarily in the nonunion sector. Since newly established firms start off as nonunion, the more dynamic the economy, the faster unions have to organize new plants simply to maintain a constant share of employment.

In the union sector, one observes a slow decline in rates in industries where unions have been strong. If an existing plant is unionized, it is likely to stay so, barring a major battle to dislodge it. An existing plant does generate rents (i.e., revenue greater than variable--but not necessarily fixed--costs) over which the parties can bargain without the plant being closed, and the workings of labor law make it difficult for management to

move the capital without first bargaining over the topic and possibly taking a strike. But little in the way of new employment is created in these plants and ultimately they are closed.

Similarly, a more mobile population is itself a result of the fact that the economy is competitive with considerable job turnover as new firms develop and old ones lose market share. Finally, a better educated population makes the supply side of the labor market more informed. This makes it difficult for firms to establish monopsonistic niches that can rely on the ignorance of the labor force of prevailing wages, benefits, and working conditions.

A fourth factor is that management opposition to unions' organizing efforts has been an important factor in explaining the difficulties unions have faced in meeting the requirements to hold a certification election and the lack of success in winning the elections that are held (Kleiner, 2002). Here again, although management opposition may be somewhat of an exogenous factor, it also ties closely with the theme of the effects of increased competition. In a competitive economy, higher wages and benefits and a loss of flexibility as a consequence of work-rule rigidities in collective bargaining agreement should be expected to generate management opposition. Similarly, the claim that management opposition is higher today than in the past is predicted by the fact that the wage premium remains high and product markets are more competitive.

The issue becomes more complex to the extent that management opposition includes an increased use of unfair labor practices. From a purely economics perspective, one can attribute the growth in unfair labor practice claims against management to a similarly rational cost-benefit calculation on the part of management. Still, the issue is highly controversial, largely because of the normative questions it poses.

A final factor involves the failure of unions to organize enough workers to offset the job losses that normally occur in a dynamic economy. Given the dynamic nature of the economy, jobs are being created and destroyed at a high rate. With more jobs disappearing than being created in already unionized firms, unions must be able to organize large numbers of new workers simply to retain even existing levels of employment. This they have been unable to do. On a purely accounting basis there are too few elections, in too many small units, and the union win rate is too low. With little in the works indicating a much higher level of successful union elections, the percentage of union employment in the private sector, on this purely accounting explanation, is likely to continue to decline (Farber and Western, 2002).

The above analysis focuses on economic factors and holds constant the underlying institutional features of the economy. Clearly, changes in the latter are also important, particularly the legal rules that structure the unionization and bargaining processes, the historical factors such as the confluence of forces in the 1930s, and the mission or function of unions.

II. Historical and Institutional Factors—Government Policy Toward Competition

The above section focuses on the prolonged decline in unions over the last several decades. To understand whether unions are likely to see renewed growth in the future, it is at least as important to understand the reasons for the rise in union density from the 1930s through the early 1950s.

In a fascinating paper, Kaufman (2002) discusses the developments of the 1930s by going back to the AEA presidential speech given by George Barnett in 1932. In that speech Barnett predicted that “unionism is likely to be a declining influence in determining conditions of labor.” The reasons cited by Barnett include reasons given today, primarily a shift in the composition of demand toward sectors in which unions have been largely unsuccessful in organizing workers and a steady increase in real wages. Of course, Barnett’s prediction was wrong for the time, in the sense that the steep upswing was about to begin. Kaufman’s punch line is that Barnett got it right, he was just forty years too early. Indeed, Barnett himself would have gotten it right if he had only made his prediction conditional on his theory. Barnett believed that the steady rise in real wages had eroded workers’ interest in unions during the 1920s, but real wages were already beginning their depression decline in 1932—hence he would have made a conditional prediction that membership would indeed begin to rise.

One of the many virtues of the Kaufman paper is to encourage humility in those who write about the decline of unionism today—believing it to be a single event rather than possibly a cyclical phenomenon. Kaufman also should remind us that hand wringing over the decline in unionism is not predicated on the unions’ fall from being a mass movement in the 1950s. When Barnett was wringing his hands, union density was at a peak of less than 20 percent of employment.

Another virtue of the Kaufman article is to remind us that reading what the ancients actually thought in the 1930s might be useful. With this as incentive, I returned to some of the older texts, including Barnett’s. But, having done so I want to quarrel with Barnett on two points.

First, even though Barnett's prediction ultimately was realized, it was for different reasons than those he outlined. But his errors are instructive. For example, the compositional shift emphasized by Barnett was the trend in employment away from craft workers, a group that historically had the highest rates of unionization. Predicting future developments based on compositional shifts is highly informative, but only if the economy is not undergoing important shifts. Barnett entirely failed to see the emerging unionization of industrial workers, a group that remains relatively highly unionized to this day. Similarly, a compositional shift analysis would have entirely missed the unionization of government workers beginning in the 1960s.⁸

More important, I believe that there is an alternative message in the historical data. The alternative message is that the sharp upswing in union density from the 1930s

to 1953 was itself highly dependent on a number of historical events coming together. While Barnett and his emphasis on the real wage decline would stress the Great Depression and the resulting passage of the NLRA as the causal factors, I would emphasize less the Great Depression and more the government's response to the Great Depression: namely a series of legislative measures that severely restricted competition.

It was widely accepted at the time that government policy during the Great Depression was anchored by two formidable trends: first, protection against foreign competition through measures such as the Smoot-Hawley tariffs (1930), and second, protection against competition at home appearing under the euphemism "fair trade practices." The National Industrial Recovery Act (1933), the Robinson-Patman Act (1936), the Davis-Bacon Act, the Miller-Tydings Act (1937), as well as state and local price-maintenance laws, were some of the policies driving attempts to "stabilize business" through reducing competition, otherwise known as "fair trade practice."⁹ Stabilizing business through fair trade was the presumed solution to the economic ills. While such a policy might appear to be a pro-business policy to modern readers, this was very much not the case.¹⁰

The fair trade policies of the 1930s were based on a particularly dark view of the large corporations that had been developing since the early 1900s. It was a strongly held belief that these large corporations were more efficient than the existing smaller companies with which they competed and would win open competitive battles.¹¹ The large corporations were also viewed as wielding considerable political clout arising from the domination of economic markets. "Fair trade," during this period meant attempting to counterbalance the existing power of large corporations.

The National Labor Relations Act, in its original form, was a reflection of the times. At a time when business firms were seen as having considerable power, the NLRA was intended to level the playing field by giving workers some offsetting power.¹² Unquestionably, the NLRA was an enormous causal factor in the spike in union density. However, its overall impact might have been a good deal less if it had been drafted in a more competitive environment. The NLRA, as amended today and as interpreted in several decades of case law, is clearly a very different measure than the original Wagner Act. While it is less supportive of unions, it is also much more in keeping with the more competitive economy promoted by today's government policies.

Unions flourished in the years between the passage of the Wagner Act and the subsequent amendments adopted in the Taft-Hartley Act. Indeed, the rapid increases in unionization came to an end in 1948, the year that the Taft-Hartley Act was passed. After 1948, unionization rates only crept slightly higher, reaching a peak in 1953, before turning negative. The timing of the passage of Taft-Hartley is consistent with the hypothesis that legal rules matter a great deal. But my point is that legal rules have to be viewed in the broader policy context in which they are adopted. It is difficult to imagine a return to a legal regime similar to the original Wagner Act in today's more competitive environment.

The claim then is that it was less the Great Depression and the decline in real wages and more a series of legislative initiatives that reduced the degree of competition throughout the economy and that allowed for the upswing in unionization rates to the extent that occurred. The ability of the newly established industrial unions in the 1930s to bargain for higher wages, while maintaining employment levels, was at least in part a function of the noncompetitive niches in which industrial firms operated.¹³

If it were a depressed economy and declining real wages that encouraged unionization during the 1930s, one should find an increase in union density in other sustained periods of real wage decline.¹⁴ However, compare the change in union density during the 1929-1936 period with the change in union density between 1972 and 1985. Both periods were one of real wage change. Although real wages fell by a greater amount during the Great Depression, workers recouped their losses much faster during the 1930s than in the later period. Unions pursued the same policy in both periods—attempting to maintain the real wages of their members. In both periods they were successful—the substantial increase in the union premium that still exists today was created during the 1970s. The question then, is why didn't union density increase during the 1970s instead of continuing to decrease? That is, why didn't unions successfully organize as many nonunion workers in the late 1970s and early 1980s as during the 1930s? My conjecture is that whereas government policy specifically attempted and succeeded in restraining competition during the former period, United States policy was specifically pro-competitive in the latter period.

The change in policy from anti-competitive to pro-competitive itself did not take place in isolation. Viewing the broader historical and institutional context, another factor was also changing in a manner that pushed policy toward supporting a more competitive economic structure. This other element was the resolution of what had been a great debate—at least by modern standards—between economists and political theorists who saw the underlying labor market as being either competitive in nature or anti-competitive or exploitative in nature. This debate has been settled in the economics community for enough years that we easily forget its importance in the days when union density was on the rise. It is to this topic that I now turn.

III. The Great Debate on the Role of Labor Unions

During the period of increasing union density, the role of unions in the economy was perceived differently than today. To many commentators unions were necessary to correct labor market imperfections or tendencies that kept wages arbitrarily, or even inefficiently, low

The modern idea that labor markets are highly competitive was only beginning to develop during this period. It was only in 1932, the same year that Barnett was giving his presidential address, that a young Sir John Hicks published his then controversial *The Theory of Wages*. Hicks's contribution was to develop what he called the marginal productivity theory of wages; or what modern readers would recognize as contemporary

wage theory.¹⁵ Hicks recognized that wages would not always clear the labor market, but argued that equilibrating forces would work in that direction.¹⁶

Most commentators, on the other hand, believed that because of market power, firms could exploit labor. The concept of an equilibrium market-clearing wage with attractive efficiency principles was very much in question even among professional economists. These commentators also believed that a counterbalance was needed to the growing economic and political power of large corporations.

These companies were seen as having not only monopoly power in product markets, but also monopsony power in labor markets. While few economists believe that labor markets are best characterized as monopsonistic rather than competitive, the issue was very much open to debate in the 1930s and 1940s.¹⁷ In monopsonistic labor markets, market power allows firms to pay workers less than their marginal productivity. Unions, by raising wages through collective bargaining, push the economy toward a more efficient equilibrium. Indeed, employment levels rise along with wages (Shove, reprinted in Hicks, 1932, p. 258). Consequently, if one believed that labor markets were primarily monopsonistic, then labor unions could raise employment, while increasing wages. Moreover, in raising wages for its own members, unions were improving economic efficiency and, in that sense, improving overall economic welfare.

Noted labor relations experts offered other arguments. For example, John R. Commons, one of the original giants of industrial relations, argued that cutthroat competition among workers set the market wage at the wage that the cheapest worker would be willing to accept (Commons and Andrews, 1927, p. 198) and that increasing wages (particularly the minimum wage) could actually increase output because employers would insist upon it (p. 223). Any number of other theories could be cited. As long as corporations were seen as having enormous economic and political clout, there was little of today's faith that labor markets generated efficient outcomes. As with monopsony, labor unions played an important productive role in these theories as well—improving the workings of the labor market.

These institutional theories took a holistic approach to labor markets that were appealing on one level, but highly confused on another. One reason Commons had a glum picture of wages falling to the level of the cheapest worker was that he assumed high levels of unemployment would drive wages lower. But this is more a theory of business cycle dynamics than of wage setting. In any case, as a matter of business cycle dynamics, wages are downwardly sticky during periods of cyclical unemployment, which is one reason that unemployment can persist for so long, a point well understood by Hicks even in 1932. But one cannot make a positive case for a continuing institutional structure such as labor unions by arguing that they might play a positive role during recessions.¹⁸

Another early and indeed continuing criticism of competitive wage theory is that markets can settle at an equilibrium wage that represents sweatshop conditions or other distributively undesirable results. But competitive results can be distributively unfair.

The competitive theory point is that the unfair results cannot be improved upon in any Pareto sense.

A final question is whether unions have distributive benefits for other labor market participants other than their own members. Two types of effects have been recognized. “Threat effects” might lead nonunion firms to copy the wages of union firms to avoid being unionized. On the other hand, the union wage premium causes a reduction in employment in the union sector that creates a “spillover effect” that can lower wages in nonunion firms. The evidence from this literature is mixed; however, any threat effect that would raise wages for nonunion members, creating a public good for nonunion workers, appears to be minor. The general result from this literature is that from a Hicksian competitive market perspective, unions can change the distributive outcome to favor their own members but do not realize allocative improvements that might benefit workers more generally.¹⁹

A more promising approach for those who continue to argue that unions have a beneficial role in improving labor market functioning is to shift the perspective from the overall market to individual firms and the potential of individual firms to be exploitative. This can occur even if the labor market itself is generally competitive. Firms, particularly those with immigrant, unskilled, or immobile labor, could act arbitrarily to set terms and conditions of employment to be below market levels. For example, Neil Chamberlain, another one of the great figures in industrial relations, wrote in 1959: “unions’ chief contribution to their members’ welfare has been to free them from the tyranny of arbitrary decision or discriminatory action in the work place” (Chamberlain 1970, p.140). Where individual firms exercise exploitative power to set wages below existing market competitive levels, the same beneficial results emerge—unions can improve the functioning of labor markets.

In terms of timing, the great debate over the role of unions continued from the early 1900s through the period of increasing union density, and it only began to lose influence as union density began its sustained decline. Simons (1944) wrote the first major academic article arguing that unions, along with corporations, had become part of the problem of concentrated power. In 1955, John R. Hicks argued that union power had put the United States on a “labor standard.” The monetary authorities would either have to validate union-imposed wage increases by inflating the currency or would suffer the consequences of recession and unemployment. More generally, as Hicks’ original ideas about the competitiveness of labor markets took hold, union power was no longer needed to improve the functioning of labor markets.²⁰

By the late 1970s, the next act played out as the government began to deregulate many of the sectors that had previously been regulated. The core idea of the 1930s—that large firms were inherently more competitive than small ones—and would therefore need to be regulated—had long since lost its following. Empirical evidence, including the studies noted above, verified that unions were largely having the economic effect of raising wages above competitive levels. In several respects the Freeman and Medoff volume *What Do Unions Do?* (1984), is the last attempt to re-establish elements of the

former theory. Although Freeman and Medoff provide much empirical evidence that has been substantiated over time, their argument that unions are primarily capturing monopoly rents or manage to increase productivity, thus offsetting the higher rates, has not been supported in subsequent empirical studies.²¹

Why is the debate important? First, it provides a picture of the public perception of unions during the period when union density was increasing. Union density was increasing at a time when public policy arguably was based on the belief that unions not only increased the welfare of their own members, but improved the workings of the economy more generally. That is, unions were seen as providing a public good. Second, the debate influenced public policy. Labor law matters a great deal. As the writing of the neoclassical economists gained greater support, it influenced public policy toward unions, starting with the Taft-Hartley amendments and continuing through the attempt of the Dunlop Commission to propose labor law changes. Finally, the success of the Hicks' model predicted the wage premium/employment tradeoff that would emerge by the 1950s.

While the debate resolved issues involving the functioning of labor markets, it obviously could not resolve issues involving the potential for exploitative behavior by individual firms. On this point, economists like Hicks could agree with labor relations institutionalists like Commons. Competitive labor market forces are not strong enough to guarantee that workers will receive the market wage.

IV. Human Resource Management and the Nonunion Sector

As noted in the last section, even if labor markets are competitive they are not perfectly so. In the employment relationship, firm-specific training creates a gap between the value of workers to their incumbent firm and their opportunity wage or next best opportunity if searching for a new job. The gap means that there is a known potential for either firms or workers to act opportunistically if they are not constrained by either legal rules or norms of good behavior. Consequently, arbitrary and unfair employment practices by an employer can render a workplace exploitative.²²

There are two related critical differences between the union sector workplace and the nonunion, human resource management-based workplace. The first and most obvious is that the nonunion sector employment relationship is typically guided by policies set by the firm using principles of hierarchy. Although employee participation may be encouraged, there is no collective bargaining or meeting of the minds to form an agreement. In most large firms, the policies follow principles of human resource management (hereinafter HR practices).²³

The second is that the human resource policies are norm-based and consequently privately enforcing rather than contractual or legally enforcing. Even where an employment practice is explicitly set and communicated to the employees, an aggrieved employee does not have standing to seek a legal remedy. The employee who has a substantive complaint can work through the procedures set by the firm. Here again, even

if the employee believes her procedural rights have been violated on top of her substantive claim, she cannot sue the employer. In its most well known formulation—employment-at-will—it says that with respect to termination, an employer may discharge an employee for a good reason, a bad reason, or no reason at all.²⁴

The rule, however, is not the practice. In practice, employers discharge employees only for cause. Consequently, employment-at-will is best understood as a “jurisdictional boundary,” that is, the courts will not intervene in the nonunion employment relationship (Rock and Wachter, 1996).²⁵ So what makes an employer adhere to the norm or good practice even when it can act arbitrarily? One specific reason is that the firm loses more in reputational harm than it can gain by acting opportunistically.

More generally, firms have found that maintaining a fair human resource policy pays off. The employment relationship is composed of highly interactive, repetitive and connected transactions. Dissatisfied employees have considerable ability to impose sanctions on firms that they believe are acting unfairly. This can occur through quits, absenteeism, and work slowdowns when the employer is most vulnerable to such activities. In the extreme cases, it can involve actual sabotage. This makes the employment relationship self-enforcing (Wachter and Wright, 1990).

Human resource policies have developed over the past two decades to make norm-based employment relationships more sophisticated and dependable in their operation. During the 1970s, when many of these practices were being adopted, a goal of the employer may have been to avoid becoming unionized. With unionization rates as low as they are today, it is less likely that practices are being driven by developments in the union sector of the economy.

Today, accepted HR practices reflect the now widely accepted view that the ability of firms to have a competitive advantage in their product market depends on their productive use of their human resources. The result in many firms is a highly structured set of human resources policies that are intended to provide a workplace based on fair substantive policies and fair procedures. This includes a mixture of external labor market comparisons as well as internal equity. Policies deal with pay and benefits, promotion opportunities, layoffs and discharges, and work rules. In addition, procedures provide details on how the substantive policies are developed and the rules that aggrieved individuals can follow. In short, all aspects of the collective bargaining agreement in the union sector find their way into human resource policies.

A major result of the HR practices is that there is less talk of exploitation today. For example, in the recent attempt by the UAW to unionize parts suppliers in the auto industry, there have been few claims of exploitative working conditions. Instead, the main justification provided is that such steps are necessary to maintain existing union strength among the unionized United States incorporated automobile makers. This is consistent with the rent-extraction story.²⁶

A regime of privately enforced employment rules and standards is obviously very inexpensive. First, it saves on what can otherwise be high legal bills. In the transaction-intensive employment relationship, the number of interactions creates the potential for a large number of disputes. In addition, and more importantly, private enforcement avoids errors by judicial or third party enforcers. Any set of contracts or rules and standards are necessarily highly incomplete given the difficulty of having rules or standards to cover all future contingencies. Consequently, the relevant rules and the facts of any dispute may be difficult or impossible for a third party to observe or verify. (Rock and Wachter, 1996)

This is not to say HR policies work perfectly. As in the union sector where the parties are occasionally forced to rely on their economic weapons or contract rights, parties in the nonunion sector are also occasionally forced to rely on their ability to sanction bad play. In addition, not all opportunistic behavior is prevented by either the threat of the application of sanctions. Individual employers can sometimes succeed as bad players. The question is not whether individual episodes succeed, but rather whether the episodes become widespread.

The softer private sanctions that workers can offer within the nonunion employment relationships are greatly strengthened by legal rules involving specific instances where opportunistic behavior by employers may succeed. This is more likely where one-shot opportunistic gains exist that can result in one-time large redistributions or where information asymmetries exist and disclosure is difficult for employees to verify. Public laws, for example ERISA and OSHA are primary examples of this phenomenon. Bennett and Taylor (2002) make the argument that unions have been important contributors to the passage of this type of legislation and, to this extent are victims of their own successes. There is a lot to be said for this position.

Another factor working to deter opportunistic behavior in the application of HR practices is the threat of unionization. Although this is of diminishing importance as union density decreases, it is still a real threat to employers that want to remain nonunion.

The upshot is that through a combination of softer, self-enforcing sanctions and tougher, legally enforcing sanctions, the HR system in nonunion firms has developed into a system that has proved both malleable and successful when measured against any number of metrics. On the one hand, as emphasized by Bennett and Taylor (2002) it can be strengthened by public legislation when the need arises. On the other hand, as emphasized by Fiorito (2002), HR practices appear to increase workers' job satisfaction.

V. The Role of the Collective Bargaining Contract

The union sector of the economy operates very differently from the nonunion sector. Whereas the nonunion sector uses norms, or HR practices, that are privately enforced, the union sector relies on third-party enforcement of collective bargaining agreement. As a consequence, the industrial relations, or IR, contracting system has high direct costs tied to its mechanism for writing and enforcing contracts. First, the National Labor Relations Board provides an administrative law system for adjudicating statutory

rights. Decisions from this body can be appealed to the federal appellate courts. Second, the common law courts provide a mechanism for enforcing the contractual rights embodied in the collective bargaining agreement. Finally, grievance machinery, including third-party arbitrators, is normally available to handle workplace disputes.

Whereas the HR workplace is inexpensive to run, the IR workplace is expensive and the reasons are essentially the reverse of those mentioned above. First, there are the direct costs inherent in a multi-layered system. Second, there are the rigidities introduced by the contract mechanism in the context of a highly transactional intensive workplace environment. Working conditions may need to be flexible and to evolve as unanticipated future states of the world occur. By including contract terms that cover working conditions, the parties reduce the flexibility in the same and replace it with a binding grievance arbitration mechanism that can generate delay and impose rigidities that reduce efficiency (Rock and Wachter, 1996).

If employers are exploitative, then the IR contracting system is better suited to protect employees than is the norm-based HR system. Legally enforceable contracts work much better than privately enforcing norms when one or both of the parties is prone to opportunistic behavior. If labor markets are monopsonistic, the IR system, by increasing wages, will increase employment as well as wages.

The contractual system also works best, at least for employees, if the primary goal of the union is to achieve an employment package for its members that exceeds competitive levels—and which management would not be willing to grant absent the risk of a strike. Norm-based, private enforcement only works well in a highly cooperative environment and this is unlikely if the parties' contract is based on the threat of the use of economic weapons.

Is there a public purpose in today's collective bargaining apparatus or is the effect only one of redistribution in favor of union members? During the great debate of the decades around the Great Depression, collective bargaining served many public purposes, including making labor markets function more effectively and serving as a countervailing force to otherwise unchecked corporate power. Today, there are still public purposes, but they are less grand in their reach and aspirations.

Perhaps the most notable public purpose is that unions encourage nonunion firms to adopt fair employment practices that lead workers to be content in a nonunion environment. And what types of unions will best keep nonunion employers in line? Certainly a collective bargaining apparatus designed to redistribute income from profits to non-management employee wages will create a strong incentive effect for nonunion employers to eschew opportunistic policies.

VI. Conclusion

In discussing the future of private sector unions and the likelihood of a resurgence in union membership one needs to evaluate the early period of union ascendancy (1930s

through the early 1950s) as well as the past few decades when unions have been in decline. We know trends currently in place are unfavorable to unions. What conditions would be favorable? When the earlier period of union growth is studied, two factors become prominent—the competitiveness of the labor market and the ability of unions to fulfill their major goal of either extracting economic rents or remedying market failures that result in exploitative employment relationships.

The rise of unions from the 1930s through the early 1950s was due to the convergence of a number of events—an economic policy that attempted to restrict competition beginning in the 1930s, the twin beliefs that labor markets were inherently noncompetitive and/or that individual workplaces were exploitative, and low union premiums. The passage of highly favorable legislation, in the form of the Wagner Act, was a reflection of the idea that unions could actually improve the functioning of labor markets and serve as a countervailing power to big business.

Over the past several decades, union density declined because government policy became pro-competitive, it became clearer that labor markets were relatively competitive, HR practices developed that reduced the amount of opportunistic behavior of employers, and unions increased the percentage premium they enjoyed in industries where rents were available. In this environment, the public-good aspect of labor unions—their ability to improve the functioning of labor markets--was called into question. The passage of amendments to the NLRA that were unfavorable to unions was a reflection of this changed sentiment as to the public good aspect of unions as well as to the adoption of pro-competitive market policies in general.

Over the near-term, union density is likely to continue to decline for all of the reasons mentioned in the first section of this paper. Attempts to forecast beyond the near-term, however, cannot be reliably done because the critical institutional and policy factors that are central to today's outcomes may change. What I have attempted to do in this paper is to focus attention on the precise nature of the institutional and policy factors that matter. For example, if employers were to become more opportunistic in their treatment of employees, union density might rise again. The recent disclosures of corporate misconduct have not had this effect, at least to this point, perhaps because the violations have been of corporate rules and standards and the losers have been more shareholders than employees.

The same would be true if government policy were to once again become anti-competitive. Examples would be renewed tariff protection, subsidies to industries suffering from competitive pressures, or re-regulation. I do not believe that a severe recession, in itself, would have the same effect on union density as the Great Depression is thought to have had. The stagflation of the 1970s, which caused a stagnation in real wages that paralleled the Great Depression (at least in terms of length), actually resulted in a decline in union density because the United States continued and indeed extended its commitment to pro-competitive policies across markets.

The one other factor that could have a major impact on union density is whether unions themselves adopt goals more in keeping with a competitive economy. Unions today are lacking one of the major factors that benefited them during the early heady expansionary years. During that period, unions were viewed as potentially correcting market imperfections, not adding to them. But the Hicksian School won the debate. Hence, a new approach would be needed. For example, unions could trade in their “monopoly face” for a more pronounced “voice face” or even a “competitive face.”²⁷ That is, they could seek to deliver an employment relationship that works more efficiently than nonunion employment relationships. At least some movement in this direction could be easily done because the union premium remains historically very high. A greater change would involve substantially more risk, particularly if union members decided that their union was no longer increasing their wages and benefits by enough to cover their membership costs. Policies that would have unions become more competitive are explored in Estreicher’s (2002) thoughtful and cleverly done paper and in Wachter (1995).

But another approach is also possible: accept the status quo. Although many academics and some union leaders believe that low union density is a major policy problem, it does not bulk large on most political radar screens. In accepting the status quo, the union movement would abandon being a mass movement that would unionize much of the eligible workforce and instead would focus on pockets where economic rents can be redistributed (Estreicher 2002) or where exploitative firms are acting opportunistically. In keeping with this change, unions would also become more of a political force advocating legislative policies favorable to workers (Bennett and Taylor 2002).

Accepting the status quo is hardly a radical position. Union density in the private sector may be low today, but “low” by what standard? As stressed above, the union movement succeeded in becoming a major force representing nearly half of the private sector workforce only briefly and only because of the unusual confluence of policies adopted during the Great Depression. Imagine that instead of adopting a series of anti-competitive policies during the 1930s, the government battled the economic downturn with pro-competitive policies. In this scenario, labor legislation is adopted, but it is the legislation that existed by 1948, with the Taft-Hartley amendments to the NLRA. From this perspective the decline in union density should be judged against some historical average rather than a peak number. Union success would not be measured against it becoming a mass movement, but rather in its ability to represent pockets of opportunity; that is, where wage premiums are possible or where exploitative nonunion employment relations exist. In this case, the predicted union density would vary over time as government competition policies and the ability of employers to sustain enlightened HR policies changed over time.

REFERENCES

- Bean, Jonathan J. *Beyond the Broker State: Federal Policies Toward Small Business, 1936-1961*. Chapel Hill: University of North Carolina Press, 1996.
- Bennett, James T., and Bruce E. Kaufman. "Introduction," In J. T. Bennet and B. E. Kaufman, eds. *The Future of Private Sector Unionism in the United States*. Armonk, NY: M.E. Sharpe, 2002.
- Bennett, James T., and Bruce E. Kaufman. "Conclusion: The Future of Private Unionism in the U.S.—Assessment and Forecast," In J. T. Bennet and B. E. Kaufman, eds. *The Future of Private Sector Unionism in the United States*. Armonk, NY: M. E. Sharpe, 2002.
- Bennett, James T., and Jason E. Taylor. "Labor Unions: Victims of Their Own Political Success?," In J. T. Bennet and B. E. Kaufman, eds. *The Future of Private Sector Unionism in the United States*. Armonk, NY: M.E. Sharpe, 2002.
- Berle, Adolf A., Jr., and Gardiner C. Means. *The Modern Corporation and Private Property*. New York: The MacMillan Company, 1939.
- Bork, Robert H. *The Antitrust Paradox: A Policy at War with Itself*. New York: Basic Books, 1978.
- Brady, Robert A. *Business as a System of Power*. New York: Columbia University Press, 1943.
- Brinkley Alan. *The End of Reform*. New York: Alfred A. Knopf, 1995.
- Joann Muller. "Autos: A New Industry," *Business Week* (July 15, 2002).
- Chamberlain, Neil C. "The Corporation and the Trade Union," In Edward S. Mason, ed. *The Corporation in Modern Society*. New York: Atheneum, 1970: 122-40.
- Commons, John R., and John B. Andrews. *Principles of Labor Legislation*, revised edition. New York: Harper & Brothers, 1927.
- Clark, Kim B. "Unionization and Firm Performance: The Impact on Profits, Growth, and Productivity," *74 American Economic Review* (December 1984): 893-919.
- Cornfield, Daniel B. "Shifts in Public Approval of Labor Unions in the United States, 1936-1999," *Gallup Poll Analyses—Guest Scholar Poll Review*.
- Dowd, Douglas F. "Against Decadence: The Work of Robert A. Brady (1901-63)." *Journal of Economic Issues*, 28 (December 1994): 1031-61.

Ehrenberg, Ronald G., and Robert S. Smith. *Modern Labor Economics: Theory and Public Policy*, 7th ed. Reading, Mass.: Addison Wesley Longman, 2000.

Estreicher, Samuel. "Strategy for Labor." In J. T. Bennet and B. E. Kaufman, eds. *The Future of Private Sector Unionism in the United States*. Armonk, NY: M. E. Sharpe, 2002, pp. 317-29.

Farber, Henry S., and Bruce Western. "Accounting for the Decline of Unions in the Private Sector, 1973-1988," In J. T. Bennet and B. E. Kaufman, eds. *The Future of Private Sector Unionism in the United States*. Armonk, NY: M. E. Sharpe, 2002, pp. 28-58.

Fiorito, Jack, "Human Resource Management Practices and Worker Desires for Union Representation" In J. T. Bennet and B. E. Kaufman, eds. *The Future of Private Sector Unionism in the United States*. Armonk, NY: M. E. Sharpe, 2002, pp. 205-26.

Kaufman, Bruce E. "The Future of U.S. Private Sector Unionism: Did George Barnett Get It Right After All?" In J. T. Bennet and B. E. Kaufman, eds. *The Future of Private Sector Unionism in the United States*. Armonk, NY: M. E. Sharpe, 2002, pp. 330-58.

Foot, Margaret, and Caroline Hook. *Introducing Human Resource Management*, 2nd ed. Addison Wesley Longman, 1999.

Freeman, Richard B., and James L. Medoff. *What Do Unions Do?* New York: Basic Books, 1984.

Galbraith, John Kenneth. *American Capitalism: The Concept of Countervailing Power*. Boston: Houghton Mifflin, 1952.

Hicks, John R. "Inflation and the Wage-Structure," reprinted in Hicks, John R., ed. *Money, Interest, and Wages*, Vol. II. Cambridge, Mass.: Harvard University Press, 1982, pp. 193-209.

Hicks, John R., *The Theory of Wages* (1932) 2nd ed. New York: St. Martin's Press, 1963.

Hirsch, Barry T., and John T. Addison. *The Economic Analysis of Unions: New Approaches and Evidence*. Boston: Allen and Unwin, 1986.

Hirsch, Barry T., and Robert A. Connolly, "Do Unions Capture Monopoly Profits?" 41 *Industrial and Labor Relations Review*, (October 1987), pp. 118-36.

Hirsch, Barry T., *Labor Unions and the Economic Performance of U.S. Firms*. Kalamazoo, Mich.: Upjohn Institute for Employment Research, 1991.

Hirsch, Barry T., and Edward J. Schumacher. "Private Sector Union Density and the Wage Premium," In J. T. Bennet and B. E. Kaufman, eds. *The Future of Private Sector Unionism in the United States*. Armonk, NY: M. E. Sharpe, 2002, pp. 92-128

Kaufman, Bruce E. "The Future of U.S. Private Sector Unionism: Did George Barnett Get It Right After All?" In J. T. Bennet and B. E. Kaufman, eds. *The Future of Private Sector Unionism in the United States*. Armonk, NY: M. E. Sharpe, 2002, pp. 330-58.

Kaufman, Bruce E. "The Organization of Economic Activity: Insights from the Institutional Theory of John R. Commons," 1504 *Journal of Economic Behavior & Organization* (2002): 1-26.

Kleiner, Morris M. "Intensity of Management Resistance: Understanding the Decline of Unionization in the Private Sector," In J. T. Bennet and B. E. Kaufman, eds. *The Future of Private Sector Unionism in the United States*. Armonk, NY: M. E. Sharpe, 2002, pp. 292-316.

Linneman, Peter D., Michael L. Wachter, and William H. Carter. "Evaluating the Evidence on Union Employment and Wages." *Industrial and Labor Relations Review* 44 (October 1990): 34-53.

Lipset, Seymour Martin and Ivan Katchanovski. "The Future of Private Sector Unions in the U.S.," In J. T. Bennet and B. E. Kaufman, eds. *The Future of Private Sector Unionism in the United States*. Armonk, NY: M. E. Sharpe, 2002, pp. 292-316.

Rock, Edward B., and Michael L. Wachter. "The Enforceability of Norms and the Employment Relationship." *University of Pennsylvania Law Review* 144 (May 1996): 1913-52.

Rock, Edward B., and Michael L. Wachter. "Islands of Conscious Power: Law, Norms, and Self-Governing Corporation." *University of Pennsylvania Law Review* 149 (2001): 1619-1700

Schuler, Randall S., and Susan E. Jackson. *Human Resource Management: Positioning for the 21st Century*, 6th ed. Minneapolis/St. Paul: West Publishing, 1996.

Shove G. F. "Review: The Theory of Wages" (1932), reprinted in Hicks, *The Theory of Wages*, 2nd ed. New York: St. Martin's Press 1963.

Simons, Henry C., "Some Reflections on Syndicalism," *Journal of Political Economy* 52 (March 1944): 1-25.

Slichter, Sumner H., James J. Healy and E. Robert Livernash. "Disciplinary Policies and Procedures," in *The Impact of Collective Bargaining on Management* 624 (Washington D.C. The Brookings Institution. 1960).

Thieblot, Armand J. "The Fall and Future of Unionism in Construction," In J. T. Bennet and B. E. Kaufman, eds. *The Future of Private Sector Unionism in the United States*. Armonk, NY: M. E. Sharpe, 2002, pp. 149-71.

Troy, Leo. "Twilight for Organized Labor." In J. T. Bennet and B. E. Kaufman, eds. *The Future of Private Sector Unionism in the United States*. Armonk, NY: M. E. Sharpe, 2002, pp. 59-76.

Wachter, Michael L., and Randall D. Wright. "The Economics of Internal Labor Markets." *Industrial Relations* 29 (Spring 1990): 240-62.

Wachter, Michael L. "Labor Law Reform," *Industrial Relations* 34 (July 1995): 382-401.

• I am grateful to Barry Hirsch, Bruce Kaufman, and Edward Rock for helpful comments and to Bonnie Clause and Greg Duffy for research assistance.

¹ Lipset and Katchanovski (2002).

² Thieblot (2002) at p. 149.

³ See the introduction and conclusions of Bennett and Kaufman (2002), the papers by Bennett and Taylor (2002), Samuel Estreicher (2002), Bruce Kaufman (2002), and Lipset and Katchanovski (2002).

⁴ Unions have certainly been successful in improving wages and other terms and conditions of employment above what would exist without collective bargaining, but the result has been job losses. Taking "wages out of competition" presumes improving member welfare without suffering the consequences of job losses.

⁵ For an earlier review of the evidence, see Hirsch and Addison (1986). This section draws heavily on Bennett and Kaufman's volume (2002).

⁶ Troy (2002) stresses this point.

⁷ Among the first papers to make this point is Linneman, Wachter, and Carter (1990). See also Hirsch and Schumacher (2002).

⁸ I argue below that the rise of industrial unions is a consequence of the anti-competitive industrial policy of the 1930s. This created rents in the newly emerging industrial sector over which firms and the new unions could bargain.

⁹ See, for example, the then current and now classic account of Robert A Brady (1943, pp. 252-53) a noted economist of the time and participant in New Deal policies. Brady clearly identifies a primary policy goal of the 1930s to be "stabilizing business," which was understood to mean restricting competition. The same theme, identified as controlling "excessive competition" is stressed by Alan Brinkley (1995) in his more recent account of the New Deal. Brady's work is honored in an essay by Douglas Dowd (1994). The National Industrial Recovery Act was declared unconstitutional in 1936. Robinson-Patman was declared by its sponsors to be the Magna Carta of small businesses since it would protect them from competition from large, more competitive firms. Miller-Tydings allowed manufacturers to specify a minimum retail price for their products in order to protect small retailers from the price competition of the chain store. (Bean 1996).

¹⁰ The idea that the antitrust laws have been used to reduce competition is stressed by Bork (1978).

¹¹ Berle and Means (1939) extraordinarily influential text on the modern corporation states: "(T)he modern corporation may be regarded... as the dominant institution of the modern world. ... (It) has brought a concentration of economic power which can compete on equal terms with the modern state. ... The future may see the economic organism, now typified by the corporation, not only on an equal plane with the state, but possibly even superseding it as the dominant form of social organization." (pp 356-57).

¹² The idea that unions represent one of the most important countervailing powers to unbridled corporate power is made most strongly by Galbraith (1952) at pp. 121-123.

¹³ This is not to say that union density would not have increased without these historical forces. The claim is simply that the extent of the rise in union density was a function of the unique historical elements of the time. What form labor legislation might have taken, absent the supervening event of the Great Depression, is too speculative for comment.

¹⁴ Real wages fell 6.9 percent between 1929 and 1933 and had recouped the loss by 1936. Real wages fell 4.7 percent between 1972 and 1981 and did not make up the loss in 1985.

¹⁵ Hicks was 28 years old when he published *The Theory of Wages* (1932). Since he was not well established until his influential *Value and Capital*, published in 1939, his ideas spread slowly.

¹⁶ Hicks (1932) at 86.

¹⁷ See Ehrenberg and Smith (2000) as concluding that labor markets are primarily competitive today. Contrast their discussion with that of Shove (1932), during the 1930s.

¹⁸ Kaufman (2002) attempts to resuscitate Commons. He is very successful in showing that modern economists can learn a great deal from reading Commons who had a great deal of knowledge of institutional forces at work in labor markets. But Commons's theory of wages, although highly attractive was wrong and gave way to Hicks's theory.

¹⁹ See Ehrenberg and Smith (2000) for a detailed discussion of the issue.

²⁰ The paper was published in the 1955 presidential address to the Economics section of the British Association for the Advancement of Science. Although he applied his argument primarily to Great Britain, he thought it applicable to the United States as well.

²¹ See, for example, Clark (1984), Hirsch (1991) and Hirsch and Connolly (1987).

²² Slichter, Healy and Livernash, (1960) report that "[t]he origin of a union in many enterprises can be traced to a belief on the part of employees that the company had been arbitrary, discriminatory, or capricious in meting out discipline. There have been foremen who acted like little czars in administering discipline and companies that imposed the most extremely personal rules upon employees. To a great extent the development of constructive personnel practices eliminated some of the abuses of 'autocratic discipline.' But labor unions must be given much credit for initiating directly or indirectly less draconic systems of discipline" (p.624).

²³ I use the term human resource practices as it is applied in the standard texts on the subjects that are taught in business schools. See, for example, Schuler and Jackson (1996) and Foot and Hook (1999).

²⁴ See Rock and Wachter (1996) for a discussion of human resource policies, and particularly employment-at-will, as being norm based. Norms are best defined as non-legally enforceable rules and standards (Rock and Wachter, 2001).

²⁵ There are obvious exceptions. The courts will intervene to enforce legislative or administrative public policies. Examples are equal pay provisions and occupational health and safety rulings.

²⁶ Reported in *Business Week* (Muller 2002).

²⁷ Using Freeman and Medoff's (1984) highly appropriate metaphor.