Working Paper No. 339

The Regulation of Collective Bargaining in Developed Economies: Lessons for Developing Economies

by

John Pencavel

July 2006
The Regulation of Collective Bargaining in Developed Economies: Lessons for Developing Economies

John Pencavel*

I. Introduction

In designing public policy towards unionism and collective bargaining, what guidance may developing economies take from the experience of the developed economies? This paper takes up this question first by outlining the activities of unions and distinguishing those people who benefit from those who lose from these activities. On the basis of this analysis, the paper proposes public policy regulations that encourage the desirable activities of unions and discourage the undesirable. At the outset, any policy prescription for developing countries must recognize a key feature of the labor markets of such countries, namely, the importance of the “informal” or “unorganized” sector in such economies.

The informal sector is defined in different ways.\(^1\) It usually embraces self-employed and unpaid family workers. To this group, the employees of small establishments are often added as are casual workers especially in rural areas. The key feature of the informal sector is really an analytical one: it consists of those workers who are outside meaningful statutory regulation and who are subject to the inexorable forces of the market. Workers in the informal sector cannot be effectively unionized. Unions are the agents of employees, not of the self-employed, and, even for employees in the informal sector, their unsteady or tenuous employment makes their organization in a union infeasible for all but a short time. This means that the sort of collective bargaining structure that characterizes some European countries such as Austria, Sweden, and Norway where the wages of

---

* This paper builds on a presentation at the conference on “Employment Friendly Labor Policies for Economic Recovery” in Jakarta on 27-28 March 2002. I thank Chris Manning and other conference participants for comments on that presentation.

\(^1\) The Report (2002) of India’s National Commission on Labour devotes much of a long chapter to the definition and vivid description of the unorganized sector. It calculates that this sector accounts for more than 92 percent of India’s work force.
virtually all workers are directly touched by collective bargaining will be impossible in developing economies. Hence any prescription for developing countries must acknowledge the existence of an important segment of the work force that is incapable of lasting organization by independent unions.

Indeed, inspection of unionized workers in developing countries substantiates this: union workers tend to be urban employees of large private firms and employees of the government. They are sometimes referred to as the elite of the working class. Rarely are more than one-fifth of all workers in developing economies union members and, typically, the proportion is much less than this.

This paper proceeds by identifying three classes of activities associated with outline. One is their role in raising the wages and other forms of compensation of their members. A second is their role as a pressure group on government. In these two roles, the union operates as a rent-seeking organization and these activities will not be interpreted sympathetically in this paper. This is because, for the most part, the benefits unions secure for their members come at the expense of the incomes of consumers and of other, often lower-paid, generate.

A more favorable view of the labor union is associated with a third activity, namely, when the union serves as the vehicle for employee participation in the workplace and, in effect, joins with management in regulating the employment relationship. In other words, employment contracts are manifestly incomplete and can be manipulated to benefit the employer at the expense of workers. The union can play a constructive role as the representative of workers. As a mechanism for employees to participate in shaping their work environment, the union has the potential to generate productivity benefits as well as further civic skills of involvement.

After outlining these activities, I sketch the appropriate public policy with respect to unionism. This policy is designed to discourage their rent-seeking activities, but to support unions as agents of employees at their place of work. The following analysis draws upon conventional economic reasoning and research. Though I am well aware of the extensive literature by legal scholars on unionism and collective bargaining, this paper does not specify in precise detail what
the law should or should not say. My purpose is to present an economic perspective and to propose a general stance for public policy, a posture, on these issues.

II. Unions in their Wage-Making Activities

Union Wage-Push

In threatening (implicitly or explicitly) collective action by employees - in other words, in collective bargaining - unions often raise wages (and other components of total compensation) above the level that would obtain in the absence of such bargaining. By how much have unions pushed wages above levels that would obtain in the absence of collective bargaining? The difference between union wages and the wages paid to the same workers in the absence of the union is called the “wage gain” of unionism. The union wage gain is not straightforward to calculate because the wages of union workers in the absence of unionism are intrinsically unobserved.

In some economies, the wage gain is computed by comparing the wages paid to union workers with the wages paid to nonunion workers. This difference is called the “wage gap”. The problem with using the wage gap to measure the wage gain of unionism derives from the fact that the wages of nonunion workers may well be affected by the presence of unions. There are several reasons for this. First, some countries have “wage extension” rules whereby the wages negotiated by unions are extended to all workers in the same industry or sector. Such rules are common in a number of Continental European countries where such wage extension rules imply that the coverage of collective bargaining agreements is very much greater than the extent of union membership. Also, for many decades last century in Australia and New Zealand, the wage “awards” issued by the labor courts were usually extended to many workers in firms beyond those expressly represented in the arbitration cases.2

Wage extension rules promote the formation and growth of employer confederations. An employer has the choice of remaining outside the organization and having his wage scales set for

2 This system of arbitration was dismantled in New Zealand by the Employment Contracts Act of 1991. Australia’s system has also been modified in recent decades.
At the same time, extension rules discourage individual workers from becoming members of unions. If a worker’s compensation is going to be the same regardless of his membership in the union, he might as well save on union dues and remain outside the organization. Hence, in France, less than ten percent of workers are members of unions and yet collectively-bargained wages cover well over 90 percent of wage and salary workers.

Union threat effects in the United States were probably greatest during and immediately after the Second World War when union membership was growing rapidly and unions demonstrated their capacity to organize large groups of workers. Sure enough, union-nonunion wage gaps were least after the Second World War (see Pencavel and Hartsog (1984)) and the threat effect is one likely explanation for this. Even today with unionism in decline in the United States, the wages of nonunion workers are higher (other things equal) in industries with an important union presence, something that has been attributed to the threat effect of union activity. However, after a careful analysis, Farber (2005) finds quite modest union threat effects in U.S. labor markets in recent decades.

Absent the threat effect, if workers are disemployed from unionized firms because of wage...

---

3 At the same time, extension rules discourage individual workers from becoming members of unions. If a worker’s compensation is going to be the same regardless of his membership in the union, he might as well save on union dues and remain outside the organization. Hence, in France, less than ten percent of workers are members of unions and yet collectively-bargained wages cover well over 90 percent of wage and salary workers.

4 Union threat effects in the United States were probably greatest during and immediately after the Second World War when union membership was growing rapidly and unions demonstrated their capacity to organize large groups of workers. Sure enough, union-nonunion wage gaps were least after the Second World War (see Pencavel and Hartsog (1984)) and the threat effect is one likely explanation for this. Even today with unionism in decline in the United States, the wages of nonunion workers are higher (other things equal) in industries with an important union presence, something that has been attributed to the threat effect of union activity. However, after a careful analysis, Farber (2005) finds quite modest union threat effects in U.S. labor markets in recent decades.
push by unions, some of these workers will seek employment in the non-union sector of the economy. The rightward shift of labor supply curves in the non-union sector tends to reduce the wages of non-union workers so that, again, the wages of nonunion workers are affected by the presence of unions elsewhere in the economy. In this instance, the wages of nonunion workers are below those that would obtain in the absence of unionism.

For these reasons - wage extension rules, the threat effect of unionism, and the wage consequences of the enhanced labor supply in the non-union sector - the union wage gain differs from the union wage gap. These reasons for the divergence between the union wage gain and the union wage gap are supported by the union wage gaps estimated in various economies. Union wage gaps tend to be largest in the United States and Canada where there are few wage extension rules and where, at least in recent years, union threat effects have been modest. On average, union-nonunion wage gaps in the United States and Canada have been around 15 percent. In Britain and Australia, where wage extension rules are more common than in North America but less common than in Continental Europe, union-nonunion wage gaps have been around 10 percent. On the European Continent, union-nonunion wage gaps have usually been estimated to be small. But this does not mean that unionism and collective bargaining has had no effect on wages in Continental Europe; it means that collective bargaining has affected nonunion wages about the same as union wages.5

The estimates of the union-nonunion wage gap in the previous paragraph are the average across all workers in an economy. There are systematic differences in the wage gaps for different classes of workers. Thus they tend to be larger for low skill workers than for high skill workers. For instance, Card (2001) estimates union-nonunion wage gaps for the United States in 1993 of about 30 percent for low-skill working men and women, but about zero for men and twelve percent

5 These estimates of union-nonunion wage gaps are derived from a number of sources including Blanchflower and Bryson (2003), Grant, Swidinsky, and Vanderkamp (1987), Hirsch and Schumacher (2000), Kahn (1998), and Kornfeld (1993).
for women at the top of the skill distribution. Similarly, using micro-data across countries, Kahn (2000) finds greater coverage of union contracts is associated with higher pay for low-skill workers relative to medium-skill workers. Schultz and Mwabe (1998) compute a union-nonunion wage gap in South Africa of about zero for highly-paid African men but of 41 percent at the tenth wage percentile. Not only are union-nonunion wage gaps larger among unskilled than skilled workers, but also unions appear to attenuate the sensitivity of wages to other characteristics. For example, the wage-age profile tends to be flatter and the wage premium attached to more schooling is lower among union workers than among non-union workers.

More generally, earnings inequality tends to be less in developed economies where unions are strong. (For instance, see Blau and Kahn (1996).) Perhaps this reflects a real causal influence of unionism: by collective bargaining and political pressure, unions push up the wages of low skilled workers more than those of high skilled workers and, thereby, they reduce wage dispersion. However, another explanation is that a society’s underlying preferences, attitudes, and conventions are reflected both in the extent of unionism and in the degree of wage dispersion. That is, countries with values that are strongly averse to wage inequality also have values that are sympathetic to union organization so that both earnings inequality and unionism are driven by the same underlying set of preferences. Obviously, it is difficult to be sure of the causal mechanism here: if public policy in a given country were altered to foster unionism, would wage inequality decline as a result? From the strength of the time-series and cross-section correlation between unionism and wage inequality in developed economies, it is difficult to resist the judgment that a policy change of this sort is likely

---

6 Card’s (2001) estimates are “raw” (that is, hold constant no other variables). Kahn’s (2000) estimates hold constant gender, schooling, age, marital status, and part-time work. Schultz and Mwabe’s (1998) estimates hold constant schooling, age, rural-urban residence, and major industry.

to reduce wage inequality. However, a policy change of this sort may well also affect the level of employment and work hours.

**Employment**

Conventional economic analysis suggests that, if the union imposes a higher wage on a single firm in a competitive industry and there is no offsetting reduction in costs, this firm will lose business, employment in this firm will shrink, and ultimately the firm will close. Examples of such union-induced closures exist and they are well-known to the union movement. This process whereby unionized firms in incompletely-unionized competitive industries shrink and ultimately close may take years to materialize, but this process means that unions must continually organize new workers to prevent a decline in the fraction of workers who are unionized. This is why unions care deeply about the rules regulating the organization of new firms: when new organization is costly, it is difficult for unionism to remain a presence in a competitive economy.

The difficulties unions face in organizing single firms in competitive industries explains why unions prefer to organize either entire competitive industries or firms with market power. When single firms are organized, they must possess some rents for unions to succeed in their wage-making activities. In effect, the union is seeking to capture a portion of these rents. But when the union captures some of the rents in this way and causes a drop in the return to capital, new non-union firms may see an opportunity to enter the industry and make a larger return. Hence unions prefer markets where entry and exit are costly or regulated by the officials of the state. If the union is well-organized, it may be able to “capture” the regulatory authority of the state and use regulation to

---

8 Of course, efforts to reduce wage inequality and support unionization may not be costless. Indeed, economies such as Germany and France where union wage gaps are small or non-existent and where the reach of unions is considerable have also been those economies in recent years where unemployment rates have been stubbornly high and employment-population ratios relatively low.

9 Inputs that are complements to union labor will also suffer from a reduction in employment. Thus, the employment of supervisory management also seems to be lower when unionism is more extensive and where the union-nonunion wage gap is greater. See DiNardo, Hallock, and Pischke (2000).
The trucking industry in the United States provides a good example of this. The union (the Teamsters) and the employers together used the regulatory process to serve their goals at the expense of consumers. Deregulation of the industry in the late 1970s caused a large drop in union wages and in the fraction of truck drivers who were unionized. Rose (1987) estimates the union was the major beneficiary of trucking regulation capturing some 65-76 percent of the total rents of the industry. Also see Hirsch (1988). Even though no explicit reference was made to wages and employment in the act that eased the regulation of the U.S. railroad industry, both wages and employment fell in the industry which was almost totally unionized. Again the suggestion is that the railroad unions were a principal beneficiary of regulation. See MacDonald and Cavalluzzo (1996). In these instances, unions able to secure the power of the state to extend their wage-making influence in much the same way as other pressure groups lobby for influence with government. A well-financed union movement is able to provide largesse to a political party and is also capable of delivering the votes of union members and the organizational structure to help legitimize the political party’s positions.

---

10 The trucking industry in the United States provides a good example of this. The union (the Teamsters) and the employers together used the regulatory process to serve their goals at the expense of consumers. Deregulation of the industry in the late 1970s caused a large drop in union wages and in the fraction of truck drivers who were unionized. Rose (1987) estimates the union was the major beneficiary of trucking regulation capturing some 65-76 percent of the total rents of the industry. Also see Hirsch (1988). Even though no explicit reference was made to wages and employment in the act that eased the regulation of the U.S. railroad industry, both wages and employment fell in the industry which was almost totally unionized. Again the suggestion is that the railroad unions were a principal beneficiary of regulation. See MacDonald and Cavalluzzo (1996).

11 For example, until the 1990s, in Britain a number of industry-specific Wage Councils operated where union representation was low. The purpose of the Wage Councils was to set minimum wages throughout the industry. Union officials often sat on the Wage Councils and the minimum wages tended to be adjusted upwards after union wages went up. This had the effect of preventing the appearance of a large gap between union wages and the wages of nonunion workers in Wage Council industries. These Wage Councils had the strong support of the union movement. Another example is provided by the “prevailing wage” laws in U.S. states in the construction industry. These laws require private contractors on government projects to pay higher wages than they would otherwise pay. The elimination of these laws in some states between 1979 and 1988 reduced the union-nonunion wage gap by about ten percentage points. See Kessler and Katz (2001).
Product market competition or the threat of competition is usually least when the employer is the state and it is a regularity of developed economies that, where unionism is not outlawed, it is more extensive in the public than in the private sector of an economy.\textsuperscript{12} This pattern is observed in developing countries, too.\textsuperscript{13} It is not merely the case that product markets tend to be less competitive when production takes place within the public sector, but also the state sometimes feels it should provide an example to private industry of what constitutes a “good” employer and this may mean being supportive of union organization.

The notion that unions find competitive product markets a less hospitable environment for their wage-push activities is supported by empirical research on union-nonunion wage gaps in both the United States and Britain. For instance, in Britain, Stewart (1990) finds union-nonunion wage differentials for most workers in competitive industries to be zero whereas they are between 8 and 19 percent in firms with some product market power. Union-nonunion wage gaps are absent in firms that operate in international markets.\textsuperscript{14} Qualitatively similar results have been derived for United States manufacturing industries. (See Mishel (1986).)

Gainers and Losers

When unions engineer higher wages in some firms, who gains and who loses from this? Clearly those workers employed at the higher wages gain at least for their duration of employment at this wage. However, in the absence of offsetting cost-reducing activities, employment will not

\textsuperscript{12} At the end of the 1980s, averaged over eighteen developed countries, the fraction of workers who were members of unions was 41 percent in the private sector and 61 percent in the public sector. See OECD (1991).

\textsuperscript{13} Thus, in Malaysia in 1985, the fraction of workers unionized in the private sector was between 7 and 10 percent whereas the corresponding fraction in the public sector was about 43 percent. See Arudsothy and Littler (1993).

\textsuperscript{14} Stewart’s sample consists of semi-skilled manual workers in about 500 private establishments observed in the 1984 Workplace Industrial Relations Survey.
expand in these firms and, more likely, it will fall.\textsuperscript{15}

We have already noted that, when unions push up wages for unionized workers, the wages of nonunion workers may rise or fall: there are wage extension rules and threat effects suggesting that the wages of these other workers will rise and there are labor supply effects suggesting that the wages of nonunion workers will fall. However, not all workers can gain: if the size of the pie is fixed, one group of workers can increase their incomes only if the incomes of another group of workers falls.\textsuperscript{16} Some workers gain from the wage-making activities of unions and some workers lose: the gainers are those who are employed at collectively-bargained wages and those who are employed in jobs where the threat of successful union organizing is real; the losers are non-union workers.\textsuperscript{17}

This is not how some observers characterize union wage-making activity. They portray union wage-making activity as a struggle over the division of total income between wages and salaries on the one hand and profits on the other hand. There may well be cases in which a union raises the pay of its members by channeling income that would otherwise be distributed in the form of profits or dividends. However, this sort of income redistribution can only be temporary and

\textsuperscript{15} I assume that union-management bargaining results in contracts where wages and employment are on the employer’s labor demand function. There is a rich literature in labor economics examining other bargaining outcomes including so-called “efficient contracts” that lie on the parties’ contract curve (which, at each wage, may be associated with higher employment than that corresponding to the employer’s labor demand function). Surely there are contracts off the employer’s labor demand function, but these are not typical because most collective bargaining contracts grant management considerable latitude in determining employment. Even if the employer is forced off the labor demand function, the reduction in profits will ultimately cause a reduction in employment below what would exist in the absence of unionism. These issues are discussed more fully in Pencavel (1991).

\textsuperscript{16} Will the size of the pie remain unchanged? We discuss below some productivity effects of unionism, but it suffices to note here that these productivity effects are sometimes positive and sometimes negative. On balance, it seems appropriate to assume the size of the pie is given.

\textsuperscript{17} There may be workers who leave the labor market altogether because they are rationed out of the union sector and yet their reservation wages are higher than the wages paid in the non-union sector.
cannot be typical. There are two reasons for this. One reason is because the profits of incorporated firms in most economies represent a small fraction of total income and when these profits are distributed over all the workers it constitutes a relatively small fillip to wages.

Consider this example for the United States whose data on corporate profits are excellent. If all pre-tax corporate profits in 1999 were confiscated and distributed equally among all employees in the economy, the average annual compensation per employee would rise by 14.6 percent. This would be a once-and-for-all boost that would need to be repeated the following year to preserve the wage increase. This number varies from country to country, but it is rarely much larger than this value for the United States in 1999. This wage increase from confiscating all corporate profits is considerably smaller than the popular impression of what would be feasible from such a transfer. The potential to transfer corporate profits to workers in this manner is constrained by two considerations: first, most incomes in societies take the form of payments to labor, not payments to “passive” agents in the form of profits, interest, and rent; and, second, any confiscated profits would have to be spread among a very large number of workers. As a result, for the economy as a whole, union-firm bargaining is less of a contest between labor and capital over the partition of 

18 In calculating this transfer, I omit the self-employed and their incomes so they are left unaffected. This is a transfer from the profits of corporations to all employees leaving the self-employed neither better off nor worse off. This estimate of the value of this transfer is very approximate as the accounting issues in defining profits and employee compensation are complex. For instance, with the growth of pension funds and employee stock ownership plans and with the use of salaries to compensate top company executives, capital and labor do not separate cleanly into well-defined groups. Nevertheless, the general sense of the per employee value of corporate profits is indicated by this exercise. This is an accounting exercise. In practice, a confiscatory policy of this sort will set in train offsetting behavior by the owners of corporations to protect their incomes.

19 Thus, in an earlier paper (Pencavel (1997)), I calculated the earnings increase from this exercise for Peru as 11 percent in 1980 and for Colombia as 15 percent in 1985. Because of the cyclical sensitivity of corporate profits, the amount of the earnings increase from this experimental transfer tends to be larger in a boom (as was the year 1999 in the United States) and smaller in a recession.

20 Thus employee compensation consists usually of between two-thirds and three-quarters of national income. The small movements in labor’s share of national income over time in the presence of drastic changes in technology and institutional structure have been a frequent topic for research.
national income and more of a struggle among various groups of workers as a whole plus those people receiving transfer income from the state.

The second reason why wage increases secured by collective bargaining are not enjoyed principally at the expense of profits is that firms are not passive in responding to attempts to appropriate their surpluses. On the contrary, firms will attempt to recover any loss of profits through union-activity by passing increased costs on to consumers in the form of higher prices.21 The prices of commodities produced by unionized firms will tend to rise and the consumers of those products will experience a reduction in their real income. The distributional consequences of these relative price changes have not been examined although they may be important. If unions tend to organize firms producing commodities that figure heavily in the expenditures of rich households, then the price effects will tend to borne more by well-off households; if unions tend to organize firms producing goods that figure more in the budgets of poor households, then the distributional impacts of relative price changes are regressive. These consumer price effects have not been the subject of scholarly research and it is not easy to assess their distributional importance.22

A firm responds to a union wage push also by changing the type of workers it employs. If the firm pays higher wages, it is likely to alter its hiring policies so as to employ people whose productivity comes closer to the higher wages. When management upgrades the quality of the work force in this manner, the union is affecting not merely the wages of workers but also the type of workers hired. Indeed, if a firm can upgrade its work force completely to match the union-imposed

21 Indeed, the drop in product demand resulting from increased prices is a principal source for the decline in employment induced by a union’s wage push. Because the employer of union workers tries to pass along union wage increases to consumers in the form of higher prices and to nonunion workers in the form of lower wages, one might envisage this situation as one in which union workers and the employer of union workers together are imposing costs on consumers and nonunion workers.

22 Part of the problem for researchers is that many commodities are not sold to households, but they are sold to other firms as intermediate goods. Thus a union-organized steel plant will pass on its higher labor costs in the form of higher prices to firms purchasing steel. How this will ultimately affect households is difficult to assess. This does not mean the distributional effects are small, only that they are difficult to compute.
Thus, in the United States, the average “raw” union-nonunion wage gap (that is, wage differences that do not control for skill or other attributes) is about 20 percent whereas the union-nonunion wage gap falls to about 15 percent when skill characteristics are held constant. Evidence of this process is provided by the fact that union-nonunion wage gaps tend to be higher when worker quality is not held constant than wage gaps that hold constant (observed) worker quality.23

In conclusion, the impact of unions on wages and employment in an incompletely-unionized economy has a number of different dimensions. When union wage push is unaccompanied by any change in productivity, then the size of the economy’s resources is necessarily constant and hence increased incomes enjoyed by union workers must come out of the incomes of others. Although commentators sometimes assume that union wage gains are secured at the expense of profits, this is largely wishful thinking. Firms are able to protect their profits by cutting employment and passing on higher wage costs to consumers in the form of higher prices. Moreover, the share of national income received in the form of profits and dividends is small relative to the number of workers so the potential for redistribution of this sort is considerably less than is widely thought. Union wage gains will then come out of the incomes of unorganized workers and out of transfer incomes.

III. Unions as a Pressure Group on Government

The well-being of union members can be materially affected by the legislative, executive, and judicial activities of government. Consequently, labor unions serve the interests of their members by playing an active role in influencing government policy. They do this by forming

23 Thus, in the United States, the average “raw” union-nonunion wage gap (that is, wage differences that do not control for skill or other attributes) is about 20 percent whereas the union-nonunion wage gap falls to about 15 percent when skill characteristics are held constant. In South Africa, “raw” union-nonunion wage gaps among African workers in 1995 were of the order of 85 percent. After accounting for differences in the characteristics of these workers, the union-nonunion wage gap falls to about 20 percent. See Butcher and Rouse (2001). Among male production workers in South Korea, Kim (1993) reports union-nonunion monthly earnings differentials in 1988 of the order of thirty percent. Once skill is held constant, the union-nonunion differentials fall to merely 2 percent. In Mexico in 1989, “raw” union-nonunion differentials of over 30 percent fall to 14 percent once skill and other variables are held constant. See Panagides and Patrinos (1994).
alliances with political parties, by calling on their members and supporters to deliver votes for particular people or particular pieces of legislation, and by using union dues and entrance fees to campaign for various people or statutes.

**Microeconomic Policy**

As a rule, unions have been skeptical of the operation of markets and have supported measures that obstruct untrammeled market forces. As examples, unions routinely support minimum wage legislation, tariffs and quotas on foreign imports, and government regulation - if not ownership - of large industries. In each of these cases, unions are furthering their members’ immediate interests. With respect to minimum wages, union workers usually earn wages higher than the state-mandated minimum wage so a higher minimum wage reduces competition from non-union firms. With respect to foreign trade, unions seek to raise the incomes of their members by curtailing foreign competition. With respect to the state regulation or ownership of industry, unions aim to “capture” the regulators and use the regulation to raise their members’ incomes.

Of course, unions are not the only pressure group on government. Indeed, some observers espouse pressure group activity by unions to offset pressure group activity by employers. However, this argument misses the fact that, on a number of issues, the interests of unions and employers coincide. Government regulation of industry often results in the unions and the employers in the regulated industry collaborating to raise their joint incomes at the expense of the poorly represented consumer. Both employers and unions in industries subject to foreign competition petition government for protection from this competition with the goal of raising their joint incomes. Employers of union labor have testified in favor of minimum wage laws alongside unions and for the same reason: the goal is to increase the costs and decrease competition from nonunion employers and nonunion labor. In each of these instances, the interests of employers and unions are congruent: employers and unions seek to enhance their shared incomes before proceeding to split this income between themselves. This greater income comes at the expense of consumers.

**Macroeconomic Policy**
On macroeconomic issues, unions usually advocate expansionary policies whose immediate impact is to increase employment even if this is accompanied by greater inflation. The bargaining power of unions is greater when workers are relatively scarce so expansionary monetary policies that decrease unemployment give unions more leverage over management. Successful assaults on entrenched union prerogatives have often been proceeded by deflationary macroeconomic policies that undermine union bargaining power. The experience of several countries in Latin America in the last sixty years is by no means singular, but it provides a telling set of examples of what has been called the macro-economics of populism in which labor unions have played an important part.

The macro-economics of populism starts with radical policies designed to effect a substantial shift in the distribution of incomes. These policies include expansionary budgetary measures and shifts in taxation and they are often joined with populist rhetoric that raises the expectations of workers. The immediate impact is a rise in real wages and in employment. In due course, the induced balance of payments deficit causes a reduction in the exchange rate and a fall in real wages. Not merely are the initial gains in real wages erased, but real wages fall below their initial level as the country tries to restore its foreign exchange reserves. The failure of the original attempt to raise real wages induces more aggressive redistribution programs and even higher rates of inflation that

---

24 Some argue that, when virtually all workers’ wages are touched by collective bargaining agreements and when the level of coordination of agreements among unions and employers is high, the price increases that result from economy-wide wage-push is so manifest that unions pull back and adopt a more moderate stance. The experiences of Austria, Norway, and Sweden are sometimes cited. The evidence from developed economies on this is mixed. Regardless, such evidence is of moot relevance to a developing economy with a large informal sector because the wages of workers in the informal sector will necessarily lie outside the scope of collective bargaining agreements.

25 Thus, consider the attack on what were called labor union privileges by Margaret Thatcher in Britain in the 1980s. One reason for their success was that, before making major changes in the law on unionism and collective bargaining, a shift in macroeconomic policy resulted in the largest rise in unemployment in Britain since the 1930s. See Pencavel (2002).

26 Excellent accounts of these episodes are provided by Dornbusch and Edwards (1991) and Sachs (1990).
East Asian economies have been less prone to these populist tendencies. This may be because the income disparities in East Asia are not so large and intractable as those in Latin America. Also, in East Asia, the fear that unions constituted a Trojan horse for Communism resulted in a greater suppression of collective bargaining so the political clout of unions was considerably attenuated.

Instances of this populist cycle are provided by Juan Perón’s policies in Argentina in 1946-49, those of Salvador Allende in Chile in 1971-73, those of José Sarney in Brazil in 1986-88, and those of Alan Garcia in Peru in 1985-88. In each of these cases, the policies were supported by many labor unions even though, at the end of these cycles, workers were worse off than at the beginning. Thus in Chile in 1971 real wages rose by 17 percent before falling by 10 percent in 1972 with a further fall of 32 percent in 1973. The support of labor unions is by no means necessary for the implementation of these disastrous policies and the policies have been pursued by governments of different persuasions including both military regimes and those of the far left. As voices for groups of urban workers, unions have typically advocated these policies in misguided attempts to effect a redistribution of incomes in societies in which income inequalities are very large.27

These policies may have been manifested most vividly in Latin America, but they are evident in other countries also. For instance, in Britain in the 1970s, a left-wing Labour Government tried to effect a substantial change in the distribution of incomes, a policy urged by the militant union movement at the time. An eruption of inflation, critical balance of payments deficits, and a wave of strikes in public sector monopolies caused the policies to be abandoned and then a totally different stance was taken by the new Conservative Government in 1979. Nevertheless, for a few years, aspects of the macroeconomics of Populism were evident in Britain.

Consider also the industrial relations systems in Australia and New Zealand at the heart of which, for most of the twentieth century, were labor courts that had the authority to impose wages ("awards") and other features of the labor contract on a broad range of firms including those that were not a party to the arbitration. Though the courts were established with the express purpose of

---

27 East Asian economies have been less prone to these populist tendencies. This may be because the income disparities in East Asia are not so large and intractable as those in Latin America. Also, in East Asia, the fear that unions constituted a Trojan horse for Communism resulted in a greater suppression of collective bargaining so the political clout of unions was considerably attenuated.
eliminating strikes,\textsuperscript{28} in fact the incidence of strikes in Australia and New Zealand was no lower than in comparable countries. The system encouraged a zero-sum game approach to labor relations and stimulated not merely the growth of unionism, but also the development of confederations of employers. Even though many agreements were reached directly between unions and employers, these agreements were reached in the knowledge that, in the last resort, either party could take the dispute to arbitration. The determination of wages at above-equilibrium levels became the fixed point in the system and this had two important consequences: the tendency for higher priced imports to substitute for more expensive domestically-produced goods required a system of extensive protection from foreign trade; and the tendency for labor to price itself out of employment required both expansionary monetary and fiscal policies to keep unemployment low and foreign exchange controls, budgetary deficits, and economy-wide controls on prices, rents, and interest rates. The relative living standards of the two countries fell throughout the twentieth century. Again, with highly politicized industrial relations systems, distributional goals propelled macroeconomic policies of accommodation that harmed economic growth and caused the economies continually to flirt with inflation.\textsuperscript{29}

IV. The Employment Relationship: The Union as the Employee’s Agent

The exchange of labor services has features that are distinctive and that make the buying and selling of labor services different from the buying and selling of other things. Labor contracts are manifestly incomplete: typically a worker is delivering to firms not merely his time, skills, and effort to management, but also a subset of his liberties. As noted by Coase (1937), whereas resources are allocated outside firms by means of the price mechanism, within firms, orders replace prices as the means by which resources are allocated to one rather than another activity. Workplaces are usually

\textsuperscript{28} Indeed, strikes were illegal for some years and subject to penalties although they were rarely applied.

\textsuperscript{29} As noted earlier, these arbitration systems in New Zealand and Australia have been altered in recent decades especially in New Zealand. An analysis of Australia’s and New Zealand’s labor markets along these lines is contained in Pencavel (1999).
very hierarchical places where one group of employees follow the instructions given by another group of employees. In these circumstances, what each worker is expected to do and what he may legitimately refuse to do are not at all well-defined. In effect, the employee concedes to management the right to assign him to a set of jobs, the range of which is usually poorly delineated.

Even if the employee’s work time can be verified, there are other dimensions of the employment relationship that are not so easily observed. These include how hard the employee works, the degree to which the worker cooperates with supervisors, and the extent to which the worker aligns his efforts with management’s goals. The difficulty in precisely defining these aspects of the labor contract explains why many contracts are not written down. Indeed, in those instances in which the responsibilities of employees have been written down, workers have been able to put pressure on employers by “working to rule”, that is, by precisely following what is written down! The vagueness that characterizes most labor contracts results in disputes over whether the terms of the (usually implicit) contract have been breached. These disputes arise because the exchange of labor services provides opportunities for one side to pursue its interests at the expense of the other party’s. Hence transactions require some sort of oversight or regulation.

One very important class of regulation of labor contracts is supplied by the state’s legal system. In many cases, the state’s regulation of labor contracts is profound and extensive. Many contracts are forbidden (such as those involving slavery and child labor) while others are subject to careful delineation (such as stipulating days that a worker is not obliged to work or activities that an employee may refuse to undertake or wages that an employer must exceed). The state’s regulations are general and, necessarily, they are not tied to each firm’s particular circumstances. Hence the legal system provides only a broad framework governing the exchange of labor services. Moreover, the state’s regulations are sometimes evaded or they are cosmetic. They must be supplemented with other mechanisms to manage the exchange of labor services.

The most common of these mechanisms is custom and convention. Such customs are
inherited from previous exchanges and give rise to repeated relationships. Labor transactions have “experience characteristics” (Nelson (1970)) meaning that the value of an exchange is difficult to estimate before an individual actually works at a job. Only after some interval of actual work can both the employer and the worker assess the value of the contract. If a mutually satisfactory contract is made, it tends to endure because the cost of transacting declines with experience with the result that long-term contracts are apt to emerge. In these circumstances, the employment relationship in any period is governed by custom, that is, by the rules that regulated the exchange in previous periods. These rules may be not only unwritten but also unspoken with the result that employment relationships are often described as implicit rather than explicit contracts.

How does the labor union change this? In some circumstances, a union may not change this at all. Some unions do no more than attempt to alter wages and leave the rest of the employment contract unchanged. However, an examination of the activities of labor unions suggests that, over time, unions respond to their members’ concerns over other aspects of the employment relation and concern themselves with issues in addition to wages. Indeed, some surveys suggest that union members do not identify higher wages as the most valuable aspect of their union’s activities. Members are more often inclined to value most the protection that their union provides against arbitrary treatment by the firm’s supervisors. In many developed countries, union contracts are characterized by the setting down of rules by which a worker’s grievance against his supervisor may be addressed. In effect, these rules constitute a system of jurisprudence internal to the firm where the union acts as an employee’s agent.

Using a different language, the union provides a “voice” for workers to express their values and preferences whereas, in the absence of such an organization, workers’ preferences must be manifested through “exit”, that is, by quitting employment and seeking work elsewhere. When there are workplace public goods such as safety and complementarities among workers, a “voice”

30 Hirschman (1970) made the distinction between “exit” and “voice” and Freeman (1976) applied it to unionism.
Thus, “In the experience of Victorian employers it was the most skilled, responsible, and steady workmen who took the lead in the unions and unionism had an uplifting influence on the weaker brethren. Industrial relations were found to be at their best when strongly organized unions entered into voluntary negotiations with stable associations of employers” (Phelps Brown (1983), p. 19). More recently, describing the Kenyan labor markets after the Second World War, Collier and Lal (1986, p.164) write, “It is potentially easier to enforce a collective agreement on unskilled labour than several hundred individual ones, partly because of reduced information costs and partly because the union has a contractual continuity which the individual worker lacks and hence has more need to maintain a reputation for sticking to contracts. This is not fanciful. Recall that in Kenya labour unrest, including strikes, preceded by many years the formation of unions. The KFE [Federation of Kenya Employers] welcomed unions as a means of replacing anarchy by order.”

Thus, Edwards (1996, p.22) observes of Indonesia in the 1990s, “......in a number of these [strike] cases, public protest is the first step in the negotiation process between ‘unorganized’ workers and managers. A significant proportion of these strikes take place outside of any process of negotiation, and among unorganized workers, that is, workers who are not affiliated to the official union or to workers’ councils”.

Some of these nonunion grievance systems have been set up by managements as part of a union-avoidance strategy. Here the “threat effect” operates on governance instead of on wages.
employment relationship. It is not merely in its role in processing employees’ grievances that the union participates in managing human resources in a company. In negotiating and helping to administer rules for hiring, training, promoting, disciplining, firing, and laying off workers, a union is participating with supervisors in the management of the company.

Industrial relations scholars have conjectured that participation begets greater efficiency, that firms with unions involved in regulating labor contracts may be more efficient than firms that lack a union serving as the agent of workers. Indeed, there is now a number of empirical studies suggesting that, in some circumstances, greater worker participation enhances productivity. In effect, these studies are suggesting that the technology of production embraces the conventions and rules associated with the use of labor and that the unionized firm’s production function differs from that of the non-union firm. This has given rise to an extensive research on estimating production functions for the two types of firms. This research sometimes supports the notion that unionized firms are more efficient than non-union firms, but there are also instances of the reverse, of unionized firms being less efficient than their non-union counterparts. A general finding seems not to be justified. When a union protects or enforces inefficient work practices such as demarcation rules that impede substitution among different types of labor and capital, it reduces the productivity of the workplace.

Whether unionized establishments are more or less productive than their non-union counterparts depends on the degree to which the parties - the workers and management representing the owners - see themselves engaged in a zero-sum game. If a zero-sum game attitude prevails,

---

34 For instance, in his classic study, Sumner Slichter (1941, p. 575) explicitly put forward this conjecture.

35 See, for instance, Freeman and Dube (2000).

36 Some of this research is described and evaluated in Pencavel (1991, pp. 34-45).

37 A zero-sum game is a situation in which there is no value to cooperation. What one party gains, the other party loses and hence the sum of the two parties’ returns is zero.
There is now a large literature on the operation of performance-based compensation schemes. For instance, Jones and Kato (1995) fit production functions to input-output information on 109 Japanese manufacturing companies over the years from 1973 to 1980. The introduction of employee stock ownership plans was associated with a 4 to 5 percent positive shock to productivity with profit-sharing providing another hike.

a mentality develops of “us versus them”: it is believed that a fixed income is to be allocated between rank-and-file workers on one hand and managers and owners on the other hand. This posture is associated with antagonistic relations between labor and management. When the union and management see themselves involved in a non-zero-sum game in which cooperation augments the pie to be divided, then union-management relations are productive. Modern human relations practices are designed to foster - though do not always secure - cooperation.

One factor that can promote cooperative relations is for the firm not to rely exclusively on time payments or piece-rate methods for compensating workers, but to link a part of workers’ earnings to the firm’s sales or profits performance. In the typical company, each worker is paid in proportion to his input of time, a payment scheme that requires the use of supervisors to monitor workers to ensure that at least a minimum level of effort is supplied. With their wage rate and work hours given, employees treat their take-home pay as roughly fixed so that the reward to any greater effort on their part accrues not to them but to the firm’s owners. Whereas rank-and-file workers are inclined to work less hard, their supervisors see their task as wringing the most effort out of the workers. The grounds for conflict are manifest.

This adverse situation can be tempered if the workers’ interests are aligned more with those of the owners of capital and this can be effected by relating a part of each worker’s earnings to the firm’s performance in the form of profit-sharing or sales-related pay. This will certainly not eliminate all sources of conflict, but it has the potential of reducing discord. It will be more effective if workers are confident that the profit-sharing system is being administered fairly and there is transparency in its operation. Here, acting as an agent on behalf of workers’ interests, a labor union can be an asset in providing the sort of assurances for workers that the profit-sharing

---

38 There is now a large literature on the operation of performance-based compensation schemes. For instance, Jones and Kato (1995) fit production functions to input-output information on 109 Japanese manufacturing companies over the years from 1973 to 1980. The introduction of employee stock ownership plans was associated with a 4 to 5 percent positive shock to productivity with profit-sharing providing another hike.
Evidence that unions may serve as the worker’s agent is provided by the role played by unionism in enforcing various statutory requirements of the workplace. As an example, consider the issue of safety. Most developed countries have regulations governing workplace safety and, if these regulations are more than cosmetic, resources must be devoted to their enforcement. Research from Britain, Canada, and the United States suggests that the level of enforcement in unionized workplaces is considerably higher than in non-union workplaces. This is the case even when, as in some provinces in Canada and some states in the United States, employee-management safety committees are mandatory in establishments. Unions tend to increase the level of awareness of safety issues in the workplace and to protect from any management retaliation employees who bring safety complaints to the notice of officials. More generally, there is evidence that unions raise the effectiveness of other statutory workplace regulations such as the receipt of workers’ disability insurance, the filing of benefits for unemployment insurance, and the advanced notice procedures of plant closings and layoffs.

In conclusion, a union alters the nature of the employment relationship. It introduces a workers’ agent into the governance of labor transactions and this may well assist the operation of statutory employment regulations. The presence of a union to serve as the employees’ agent may enhance a firm’s efficiency or to impair its efficiency. There are examples of both labor relations being better in unionized firms and of the reverse. This makes generalization impossible. Altering payment methods so that a component of workers’ pay is associated with the firm’s earnings can

39 Because profits are highly cyclical, a profit-sharing system also imparts some flexibility to workers’ earnings which confers macroeconomic advantages in times of recession.

40 See Reilly, Paci, and Holl (1995), Bernard (1995), and Weil (1999). The British study found no difference in workplace injuries between union and nonunion establishments provided employees participated in the safety committees, but workplace safety was less in establishments where safety was monitored exclusively by the representatives of management.

help in aligning the workers’ interests and the owners’ interests. It is worth emphasizing that any governance benefits from unionization accrue at the place of work, that is, at the plant or establishment. It is the union operating at the workplace that may help to improve the performance of the firm and improve the welfare of workers. It is not the national union or some union confederation that delivers these useful services.

V. Prescription

Given this analysis of the activities of labor unions, what is the appropriate posture for public policy toward unionism? The argument has proposed that it is the activities that the union engages in at the workplace in the transaction of labor services that may merit support from public policy, not the union in its wage-making role nor in its political role.

Every country confronts the challenge of designing a structure to reduce the unfavorable features of unionism and to encourage its beneficial activities. An aspect of this policy question is to determine which issues are to be governed by statute law and which issues are to be subject to the self-regulating operations of markets. And, for those issues within the domain of statute law, should the law favor one side or the other and should the regulations take a very specific form about how the parties are to behave? In open societies, in view of the vagueness of the employment relationship, all-embracing legal regulation is impossible. That is, the precise delineation of what the employer and the union may or may not do in various circumstances and in all types of employment is infeasible. On the other hand, it is also unimaginable for a society to dispense with all legal regulation of the employment relationship. The question is where to draw the line: what should the law specify and what should it be silent about.

Labor lawyers sometimes favor a broad reach for the law and advocate a legal system that actively supports unionism. For instance, the eminent British lawyer, Otto Kahn-Freund wrote, “The main object of labor law has always been, and we venture to say will always be, to be a countervailing force to counteract the inequality of bargaining power which is inherent in the employment relationship” (Kahn-Freund (1983, p. 18)). Lawyers arrive at this position by focusing
on the conflict of interests between workers and management. A conflict of interests over the division of an enterprise’s surplus certainly exists. In addition, however, there is complementarity of interests between workers and management, namely, that of maximizing this surplus before settling on its distribution. This is why labor contracts contain elements both of conflict and of harmony. If the interests of workers and management are partly in opposition and partly in accord, why should labor law should be “based on anything other than conflicting interests” (Wedderburn, (1986, p. 26))?

An economic analysis of unionism suggests three general principles to guide public policy with respect to unionism and collective bargaining. The general goal of these principles is to moderate unions in their rent-seeking activities and to bolster their workplace role in guarding the interests of employees.

Decentralization

The first of these principles is Decentralization: as far as possible, the focus of industrial relations should be at the workplace. In this way, the source of the union’s strength is close to - not remote from - the individual union member. When authority within the union resides at the local level, members have the opportunity to participate in meaningful ways to influence the policy of their organization. Their membership means more than simply paying dues to a distant organization to support individuals whom the individual union member does not know.

In addition, the union is more likely to serve as an effective agent for its members if its authority lies at the workplace. The mutual dependence of the workers’ and the capital owners’ interests is likely to be most evident when collective bargaining takes place at the level of the workplace and, consequently, productivity-increasing agreements have a greater chance of being negotiated. There is no guarantee of constructive labor relations, but at least the consequences of the union-management agreements are more apparent and they may be fashioned with attention to the distinctive idiosyncracies of the particular workplace.

Moreover, because labor demand functions tend to be more wage-elastic at the level of the
workplace than at some aggregate level such as the industry, the constraint on union wage-push provided by employment reductions (or lack of employment growth) is greater when collective bargaining takes place at the workplace.

Furthermore, the union’s rent-seeking effectiveness in influencing national economic policy is less if the core of the union’s strength is the local workplace. The locus of union authority tends to be found where collective bargaining takes place so, if collective bargaining were undertaken by national unions at the industry-wide level, their pressure group activities on government would be more influential.

The complete opposite from decentralization occurs when the union’s authority lies with confederations of unions that negotiate economy-wide wage agreements with confederations of employers. Arrangements such as these occur in some European countries such as Austria and Sweden, but the reach of such agreements can never be as extensive in developing countries where there always exists labor markets outside the scope of collective bargaining. The notion of economy-wide collective bargaining agreements is simply infeasible when there is a large informal sector.

It is possible in developing countries for collective bargaining to take place at the industry level in the subset of industries that are unionized. In this event when wage bargaining takes place at the industry-wide level instead of at the firm, union wage-push will tend to be greater, confederations of employers (another source of coercive power) will develop, an oligarchic and detached union leadership is likely to be cultivated, and an influential pressure group on national government will emerge. In this setting, unionized workers tend to constitute a select and privileged group of workers and a significant gap in earnings will arise between those fortunate enough to be covered by union agreements and those not so fortunate.

---

42 After examining unionism in six developed countries, Clegg (1976) reaffirmed, “...power within unions is concentrated at the level where collective bargaining is conducted. Industry bargaining centralizes power at union headquarters. Regional and plant bargaining decentralize power” (p. 41).
The State’s Disengagement

In some countries, the state plays a conspicuous role in overseeing labor relations. This conspicuous role takes two forms. In Obstructionist regimes, the state actively discourages, if not suppresses, independent labor unions. There are many techniques to effect this. For example, the benign-sounding rule that requires each union to register with the state may be used to determine which unions are to be permitted and which to be restrained.\(^{43}\) In such countries, union leaders may be viewed as representing a political challenge to the incumbent government which may use heavy-handed techniques of intimidation to deal with this threat. Strikes tend to have serious political overtones and are viewed as potential threats to the authority of the government.\(^{44}\)

The other type of regime where the state occupies a conspicuous role occurs when collective bargaining is decisively sponsored as the appropriate means for determining wages. In these Patronage regimes, there are often formal links between one or more of the major political parties and the labor unions\(^ {45}\) and the state plays an active part in preventing market forces from

---

\(^{43}\) Thus, Freeman (1993) has argued that “repressive regimes [in East Asia] have used [the requirement of unions to register with the state] to direct workers into acceptable unions and to make other unions unacceptable” (p. 45).

\(^{44}\) Indonesia under Soeharto provides a useful example of an Obstructionist regime. Manning (1998, pp. 209-11) writes: “The Soeharto regime opted for a government-controlled, national trade union organization, the All-Indonesia Labour Federation, FBSI (later the SPSI) in 1973....The SPSI emerged as a tool for government control of organised labour, rather than representing workers’ interests in negotiations with employers. The leadership was weakened and its independence compromised from the early years, and it failed to promote labour issues at a national level. Leaders were replaced by appointees more sympathetic to the government and the union’s leadership was also closely vetted by the military. A large number of public sector workers, including employees in state enterprises, were excluded from the organization, weakening its influence in modern manufacturing.... [M]any enterprise unions were little more than token organizations....The SPSI’s inability to effectively defend workers in industrial disputes undermined its credibility. Local police and military were actively engaged in dispute settlement, usually at the behest of employers. Dissident workers and strike leaders were detained, interrogated and frequently subjected to physical abuse by the authorities.”

\(^{45}\) In the Indian sub-continent, each of the major political parties is allied with its own labor union confederation so the competition among political parties finds expression also in competition among
constraining the effects of union activities. This may mean requiring nonunion firms to pay union or close-to-union wages, imposing tariffs or quotas on import-competing products, obliging firms to pay workers severance pay, and preventing firms from closing plants without the prior permission of the state. There is a tendency for strikes to be viewed as inconvenient expressions of legitimate worker grievances.\footnote{46}

Patronage and Obstructionist regimes appear to be at the opposite ends of the spectrum and, indeed, they have contrasting attitudes toward collective bargaining, one being inimical to it and the other promoting it.\footnote{47} However, both regimes share the feature of the state being very intrusive in the regulation of unions and collective bargaining. When the state occupies such an interventionist role, rival union federations.

\footnote{46} Argentina provides an example of a Patronage regime and the sort of damage wrought by a highly politicized union movement. Juan Perón actively courted the unions while serving as Secretary of Labor from 1943 to 1945. When Perón was imprisoned in 1945, the unions organized a general strike to secure his release and then worked (successfully) to get Perón elected as President in 1946. During his authoritarian rule, not only were many industries nationalized and domestic industry heavily protected from foreign imports, but also the state intervened extensively in labor markets. Union membership was made compulsory in the public sector, collective bargaining was designated to take place at the industry level with wage increases applying to union and non-union workers alike, labor union leaders were appointed to many political positions, unions were granted monopolies of representation, the collection of union dues was required of employers who handed them over to union leaders, a portion of all wage increases was transferred as income to each industry’s union, and the state’s welfare systems for pensions and health were integrated with the union-run schemes and union leaders were given the authority to administer them. Though Juan Perón was overthrown in a military coup in 1955 and some of these regulations were altered, each time a quasi-Peronist administration takes office, many of these regulations reappear and serve as a drag on the performance of Argentina’s economy. The labor unions have consistently been one of the major political forces in the country. See, for instance, Epstein (1989).

\footnote{47} Some countries provide interesting mixtures of the two. For instance, consider South Korea where for many years the state has closely regulated the activities of labor unions. Occasionally, strikes were dealt with by arresting union leaders and disputes were settled by compulsory arbitration. Since 1987, labor relations have been liberalized somewhat although there remains a concern that the government will intervene in instances when it disapproves of important collective bargaining outcomes. Coupled with this heavy regulation of industrial relations is a labor market policy that actively discourages layoffs and that has encouraged the view of the state as having a heavy obligation to ensure rising real wages for workers.
posture, unionism tends to become highly politicized. It seems to workers and union leaders that their welfare is affected less by the economic realities of the workplace and more by the impulses and judgments of politicians. Because the state’s posture seems so critical in specifying their potency, the unions’ dealings with politicians are more consequential than the unions’ relations with the employers of the workers they represent. Increasingly the union leaders become detached from the specific conditions of any given enterprise and more absorbed in the resolution of issues with politicians. When labor relations and collective bargaining become divorced from the economic realities at the workplace, the tendency is for wage and other agreements to reflect political compromises rather than the economic circumstances of a particular firm.

Both Patronage and Obstructionist regimes are better avoided. Economic efficiency is promoted if, as far as possible, the state adopts a posture of disengagement from unionism and collective bargaining. The state neither promotes nor repudiates collective bargaining. Employers are neither obliged to recognize labor unions nor obliged not to deal with them. Private sector firms and unions develop their own bargaining protocols with the state neither specifying whether they should bargain and how they may bargain. The framework for unionism and collective bargaining should not show preference for either party nor encroach on their decisions except to protect the welfare of those who are not explicitly represented in the bargaining. The state should leave as much for the bargaining parties themselves to determine so that a diversity of collective bargaining protocols results; there is no attempt to fit all collective bargaining structures into one model.

This diversity will imply that, in some workplaces, one union represents all the workers while in other workplaces many unions exist. Need the multiplicity of unions be a problem? No. The existence of many associations suggests a heterogeneity of preferences and interests among workers and, where possible, it is desirable for this heterogeneity to be accommodated. The existence of many unions does not necessarily mean that an employer must negotiate separately with each union. If an employer wants to avoid a multiplicity of contracts, the unions may be required to form a single bargaining committee for the purpose of negotiating a contract. In this way, multi-
unionism is compatible with one collective bargaining agreement. This arrangement preserves the benefits of multi-unionism.  

The benefits of multi-unionism are the familiar advantages of competition - in this case, the competition for members. A single union granted protection from competition from other unions is more likely to be managed by a small, entrenched, oligarchy of leaders with interests separate from the rank-and-file members. It is not surprising that authoritarian governments tend to favor monopoly unionism (that is, permitting one and only one union to operate in a given sector) as this helps the state oversee and dominate the union movement.

The state's disengagement from collective bargaining and unionism is most difficult to effect when the state is the employer. When an industry has been nationalized, its employment and wage decisions are apt to have heavy political overtones. Dispassionate judgments about alternative ways to manage the industry become clouded with the political implications of various actions. Thus,

---

48 Will workers have more bargaining power if they organize in a single union or if they bargain separately in different unions? A general answer is not possible. For instance, suppose (as has been argued throughout this paper) that employment reductions act as a deterrent on wage-push by a union so that, in general, a union will push up wages less when the labor demand function is more wage-elastic. Then, in the case of two types of labor, say, type A and type B, the wage-elasticity of demand for type A increases (in absolute value) with the share of total costs paid to type A workers if it is easier for firms to substitute away from A workers in production than it is for consumers to substitute away from the product made by A and B labor. In other words, the wage-elasticity of demand for a given type of labor may rise or fall with the share of that labor in total costs so there is no general rule about whether the workers should split for purposes of bargaining or unite. Other models comparing bargaining outcomes by a single union with outcomes involving separate unions are found in Rosen (1970) and Horn and Wolinsky (1988). In these papers, the manner in which the types of labor figure in production turns out to be the key determinant of whether wage gains will be larger in one case or the other. Again, a general statement is not possible.

49 There are instances of a single union avoiding the creation of a privileged, firmly-ensconced, leadership. An example is provided by the old International Typographical Union in the United States in which two parties regularly competed for power and the democratic traditions resulted in the routine turnover of the leadership. See Lipset, Trow, and Coleman (1956). However, this example tends to be the exception and competition among unions (instead of within a union) is more likely to avoid the creation of an oligarchical leadership that enjoys special entitlements.
although sound management may require adopting a “tough” stance on wage increases or on layoffs, this becomes politically awkward if the industry’s management is perceived to be closely linked to the government. When faced with immoderate demands from the labor unions, the ministers with responsibility for the nationalized industry are inclined to be prodigal with the state’s coffers and, in this way, not jeopardize their political standing.

This is an argument, of course, for keeping state ownership to a minimum. In those instances where the state cannot avoid being the employer, hard budget constraints need to be enforced so that the sort of employment-wage trade-offs that face private sector unions also apply to the public sector. The management of these state industries must be headed by people whose employment is protected from the vagaries of politics.

Because the public sector tends to be monopolistic, industrial disputes in the public sector tend to impose severe costs on consumers. The services of many public sector workers such as police officers, fire fighters, hospital workers, and utility workers have few immediate substitutes so, when these services are interrupted through a strike or lockout, their loss imposes considerable costs on consumers.

In such circumstances, compulsory arbitration has much to recommend it: that is, although “essential” public service workers such as the police and fire fighters may be represented by labor unions in collective bargaining, in the last resort the workers may not strike nor the employer may lock out the workers. Compulsory arbitration may take different forms. Conventional arbitration allows arbitrators to devise any settlement from the parties’ last offers and demands and the tendency seems to be for the arbitrators to settle on some middle position between the two extremes. Knowing this, each party has the incentive to make more extreme demands and offers and the parties see little reason to avoid arbitration.

By contrast, with final-offer (or “pendulum”) arbitration, the arbitrator\(^{50}\) is required to accept

---

\(^{50}\) Although I write “the” arbitrator, typically final-offer arbitration works with several arbitrators. There are many different methods to select the arbitrators and generally the parties’ preferences are
either management’s final offer or the union’s last demand. The preference for final-offer arbitration does not arise because final-offer arbitration produces different wage decisions from conventional arbitration; on the contrary, there is little evidence that, on average, the wage decisions of the two arbitration systems are different. The appeal of final-offer arbitration is that, whereas conventional arbitration encourages the two parties to make more extreme demands in the expectation that the arbitrator will split his decision, final-offer arbitration induces the parties to come together and make compromises out of the fear that the final-offer arbitrator will impose the other party’s last offer on them. In other words, when the parties are risk averse, the concern that one party’s last offer will appeal more to the final-offer arbitrator prompts the other party to soften its position and this makes agreement between the two parties more likely. And an agreement arrived at by the two parties is more likely to be accepted by them and made to work than an agreement imposed by a third party. Thus, one might conjecture that, once the parties become familiar with its operation, a system of final-offer arbitration is one where arbitrators are called upon less frequently than under conventional arbitration.

The criteria directing the arbitrator’s choice need to be clearly stated so the parties may form better estimates of the arbitrator’s decision. These criteria include changes in the cost-of-living and

taken into account. Some of these methods are described in Lester (1984).

51 Thus, final-offer arbitration resembles many decisions in criminal cases in which a judge or jury has to decide on the guilt or innocence of a defendant. Except in some special instances, there is no half-way position.

52 There is some evidence of this from the use of final-offer arbitration in private sector bargaining in Britain where final-offer has often been an ingredient of the “new style deals”. Metcalf and Milner (1992) find that there have been fewer failures to agree (and thus fewer instances in which arbitration has been invoked) when the agreements specified final-offer arbitration (at least when such arbitration is accompanied with conciliation and mediation procedures). Note that this evidence is somewhat difficult to interpret because, in this British case, final-offer arbitration has been selected by the parties as a device to help resolve disputes. That is, instead of the association measuring the impact of final-offer arbitration on bargaining impasses, the correlation also reflects causation running in the reverse direction - from the likelihood of bargaining impasses to the selection of final-offer as the arbitration mechanism.
in productivity. The arbitrator would take into account whether it has become difficult to hire workers of sufficient quality or whether the supply of such workers has been ample.

Though this section has called for the state’s disengagement from collective bargaining as much as possible, it needs to be emphasized, as mentioned earlier, this is not an argument for the absence of law. Far from it: associations of workers in unions are a key feature of open societies and the law should ensure that individuals have the opportunity to form such groups without fear of reprisal from other organizations. However, once the broad framework establishing the fundamental right of workers to form associations free from intimidation has been established, the state’s task is to be as uninvolved in the operation of collective bargaining as possible. When the state goes beyond enforcing the rules of the game, collective bargaining is apt to become politicized with the result that wage and employment outcomes reflect less a firm’s economic conditions and more the political circumstances of the setting.

**Competitive Markets**

If collective bargaining and labor unionism are not to be regulated by detailed statute law, are they exempt from oversight? The answer is that, if permitted to function well, competitive markets will provide effective regulation of unionized markets. In other words, the extent to which the undesirable aspects of unionism can be avoided depends not merely on the regulations operating in the unionized sector of the economy, but also on how effectively other markets - nonunion labor markets, product markets, financial markets - function. Competition in other markets tends to enhance the wage-elasticity of labor demand functions in the union sector and large wage-elasticities imply sharp reductions in employment from monopolistic wage push by unions. Such sharp reductions tend to serve as a check on the extent of wage-push and thus provide a self-regulating dimension to union wage-making activity.53

---

53 Strictly speaking, this statement depends on the precise form of a union’s objectives and on bargaining protocols. However, a general sense of the relationship between union wage-push and the wage-elasticity of labor demand functions is provided by the following simple characterization. Suppose a union’s objectives (as indexed by $U$) depend on both wages ($w$) and employment ($E$) in
Hence curbs on competition - such as wage extension rules that broaden the reach of union wage-setting, support for product market monopolies, and impediments to free trade - are to be avoided. Also, for the employment implications of wage push to serve as an effective constraint on union wage practices, macroeconomic policy needs to avoid expansionary fiscal and monetary policies that conceal the employment-reduction effects from manifesting themselves. This means shunning what Hicks called the Labour Standard, that is, making wages the fixed point in the economic system and conducting monetary and fiscal policy to adjust quantities and other prices so that these wages may be rationalized.\(^{54}\)

The presence of competition will not only minimize the costs of monopolistic wage settlements, but competition will also tend to reduce the costs of arriving at these settlements. The principal costs of negotiating settlements are associated with interruptions to production caused by strikes and lockouts. When the output of a single firm in a competitive industry is halted through an industrial dispute, the costs of the strike or lockout fall on the parties to the dispute - management

\[ U = (w - v) \theta E \]

where \( v \) is the wage that would exist in the absence of the union. The union prefers both higher wages and higher employment and the trade-off between higher wages and employment depends on the value of \( \theta > 0 \). Suppose also that the labor demand function, \( E = f(w) \), acts as a constraint on the achievement of both higher wages and higher employment. If wages are set at a level that maximizes \( U \) subject to the labor demand function, then \( \Delta w = \theta/\eta \) where \( \eta \) is the elasticity of demand for employment with respect to wages (that is, \( \eta = -(dE/dw)(w/E) > 0 \)) and \( \Delta w \) measures the proportional wage push (that is, \( \Delta w = (w - v)/w \)). Clearly, \( \Delta w \) is smaller as \( \eta \) gets larger; the degree to which the union pushes up on wages falls as the elasticity of demand for employment with respect to wages gets larger. Anything that causes \( v \) to rise will also induce an increase in the union wage, \( w \). Of course, the employment-reducing impacts of union wage policies will not deter union wage push if the union does not care about employment. However, there is little evidence that employment is not a consideration in union objectives. More formal models of union objectives and bargaining may be found in Pencavel (1991).

\(^{54}\) Hicks (1955, p. 391) wrote, “....the world we now live in is one in which the monetary system has become relatively elastic, so that it can accommodate itself to changes in wages, rather than the other way about. Instead of actual wages having to adjust themselves to an equilibrium level, monetary policy adjusts to the equilibrium level of money wages so as to make it conform to the actual level. It is hardly an exaggeration to say that instead of being on a Gold Standard, we are on a Labour Standard.”
and union labor - rather than on consumers who have alternative places to shop. In such situations where the costs of the dispute do not fall on vulnerable third parties, the state may safely adopt a neutral stance. This means not providing subsidies to either labor or the firm engaged in the dispute. Just as the striking workers should be eligible for (temporary or permanent) work elsewhere so the struck firm should be free to attempt to resume production with non-striking workers.

Health and Safety Committees

The stance for public policy toward unionism sketched here is likely to be effective in constraining their rent-seeking activities, but should more be done to encourage the participation of employees in shaping their work environment? The case for greater participation rests on evidence suggesting productivity benefits to firms from engaging workers more fully in designing the workplace. In addition, there are wider benefits in inculcating the virtues of participation in civic activities. The political case for democracy is easier to make and effect if the daily experiences of people at their place of work are characterized by engagement and association. Therefore, what can be done to stimulate greater worker participation in decisions at their place of work?

Consider the notion of requiring each establishment to have a joint employee-management committee to address health and safety issues at the workplace. The purpose of such committees is to raise awareness of and to bring safety matters to the notice of both management and workers. A committee of this sort will tend to improve information about safety which has immediate benefits for employees. In addition, such committees may help raise the collective consciousness of the workers and infuse a wider civic sensibility. In other words, health and safety committees may cultivate the sort of cooperative outlook that evolves into a local union and that encourages the agency role of unionism.

This mention of health and safety committees serves as a reminder that labor unions are not

---

55 It would require a safety code to be outlined by statute that provides guidance for workplaces with an administration to administer the code. Workplace committees would have the task of implementing and enforcing the code.
the only “voice” mechanism for employees. A number of European countries have Works Councils which are organizations of worker representatives usually found in large or medium-sized establishments. The principal role of Works Councils is not to bargain over wages (indeed, the typical Works Council is not engaged in wage bargaining at all), but to serve as a device for management and workers in a given plant to discuss issues of common concern and to help ensure equitable treatment of employees.\textsuperscript{56} One important issue is whether Works Councils serve as a substitute for unionism or as a complement. In countries such as Germany where wages are usually negotiated at a level higher than the individual firm, the Works Council is very much a complement in that the Council addresses itself to workplace-specific issues of employment regulation. In other circumstances, the Works Council delivers services that the union may aspire to administer and the presence of a Works Council in a workplace will reduce the demand for union representation.

VI. Summary and Conclusion

Three classes of labor union activities have been identified. The first consists of their wage-making activities associated with collective bargaining. The important questions to ask of collective bargaining are by how much are wages (and other components of total compensation) pushed up and who bears the cost of these wage increases. The ability of unions to push up wages depends on many factors, but it is usually argued that their wage-push capacity tends to increase if the employment-reducing impacts of their activities are small. In other words, unions tend to push harder on wages when labor demand curves are relatively inelastic and employment declines are small or difficult to perceive. This is why unions tend to prefer bargaining with groups of employers for whom labor demand curves are more inelastic than they are for any single employer. Consistent with this, union-nonunion wage gaps are larger for firms with market power than for firms in product market competition.

On the issue of who bears the cost of these wage increases, it is difficult to avoid the

\textsuperscript{56} See Rogers and Streeck (1995).
conclusion that they are borne by non-union workers. Although there may be instances in which wage increases are won at the cost of a reduction in profits and dividends, this cannot be the general case. Capital is so mobile that it constantly seeks that activity yielding the greatest net return on capital invested. Attempts to redistribute income from capital to labor induce defensive reactions on the part of management, the representatives of the capital owners, to preserve their returns. These defensive reactions include cutting employment, substituting other inputs for labor, passing cost increases on to consumers in the form of higher prices, moving production to plants outside the influence of unions, and sub-contracting out to nonunion agencies certain separable activities. And any reduced returns to capital will simply discourage further investment in the unionized firm.\textsuperscript{57}

Hence, if the national income is given and capital-owners are successful in defending their incomes, higher union wages must come out of the nonunion workers and consumers. Some non-union workers may benefit from unionization in that the threat of unionization induces their employers to raise the wages of their workers to reduce the likelihood of their organizing. But, in an economy with a significant nonunion sector that is difficult for unions to organize (as is the case with most informal sector jobs in developing countries), union threat effects cannot be pervasive and it will be these vulnerable workers who will tend to bear the costs of union wage-push.

The second class of union activities relates to their actions as a pressure group on local or national government. Usually, this rent-seeking activity takes the form of seeking impediments to the operation of markets for the benefit of a segment of the work force. Unionism is by no means the only pressure group endeavoring to enlist the coercive apparatus of the state to secure prerogatives, but it can be one of the most effective. This is because it claims to speak on behalf of all workers and of the underprivileged. In fact, in most economies and in certainly the developing economies, unions represent a subset of all workers and they tend to be more in the way of the well-
paid workers than the low paid.⁵⁸ In some instances where confederations of unions have extracted prerogatives from the state, unionism represents a major political force that, at worst, may represent a threat to the ultimate authority of the state.

The third dimension of unionism consists of their agency activities at the workplace.⁵⁹ As argued above, the employment relationship is unlike the textbook contract where a well-defined commodity is exchanged on clearly-specified terms. On the contrary, the typical labor contract is incomplete and provides ample room for opportunistic behavior. Employees tend to be especially vulnerable to actions that well-informed managers can take to improve their welfare at the expense of their workers. The law tends to constrain malfeasance, but typically only broad constraints are established and the cost of enforcing the law leaves unprincipled employers with opportunities to enhance their interests at the expense of workers. In this context, a union may serve as a useful agency on behalf of workers to ensure that contracts take a less arbitrary and more balanced form. In this case, the union participates with the employer in the management of the company at least insofar as the delivery of labor services is concerned. In serving as the workers’ agent, the union can bring an adversarial posture to labor relations or it can contribute to a harmonious and

⁵⁸ Kilby (1967) analyzed Nigeria’s unions shortly after Nigeria’s independence from Britain. In his conclusion, he drew inferences from his study of Nigeria about the typical situation in developing countries. What he wrote in 1967 is applicable today: “....rather than being an exploited group, organized labor is already a highly privileged minority....[T]he [unionized] laborer’s earnings considerably exceed his opportunity income outside the organized employment market - in short he is enjoying a higher standard of living than he has ever before known.....Under existing conditions in underdeveloped countries,.....labor organization becomes a political instrumentality for channeling the protest of relatively privileged wage earners against the distribution of the national wealth between themselves and those at the apex of the distributional pyramid.” (Kilby (1967), pp.515-6.)

⁵⁹ Many early unions acted as “friendly societies” for their members providing from union funds an array of insurance benefits covering sickness, unemployment, old age, and death. These mutual insurance activities still remain in some unions and, indeed, are growing in Britain where unions offer subsidized legal services. In certain countries, unions administer the state-financed unemployment-insurance system and some have argued this encourages the unions to be concerned with the welfare of the unemployed (sometimes labelled the “outsiders”) as well as the employed (the “insiders”).
productive working relationship with management. If the union and management can move beyond a zero-sum mentality, there is the possibility that the participation of the union in labor contracts can bring benefits to workers and the firm.

In view of this assessment of unionism, an appropriate public policy would call for discouraging the rent-seeking activities of unions and encouraging their role in serving as the agent of employees and in developing social capital. Three general principles are put forward to advance these goals: the decentralization of collective bargaining; the disengagement of the state from many issues associated with unionism and collective bargaining; and the promotion of competitive markets throughout the economy. To enhance the participation of employees in shaping their work environment, the establishment of worker-management committees to address health and safety committees would not merely raise the level of safety at the workplace, but also develop aspects of civic engagement that raise the quality of social intercourse and, ultimately, representative government.

These principles complement one another. That is, the advantages to decentralization are greatest when markets are competitive: the gain from firm-level bargaining is meagre if these firms are all monopolies that can pass on wage increases to consumers. Or the benefits from the state’s disengagement from collective bargaining are undermined if union-management wage agreements apply to broad swathes of industry and consumers and informal sector workers bear the costs of this vigorous wage-push activity. Hence the returns from implementing these three principles are greatest if all are implemented together.

If these principles were adopted, it is likely that, in developing economies, only a minority of workers would be unionized as is the case at present. But in those sectors where unions were active, they would have the opportunity to share the organization’s rents with the employees. Most important, these principles would foster economic growth and increases in productivity. Ultimately, increases in the real incomes of workers are associated with increases in productivity so policies that are conducive to the growth in productivity are, in the final resort, pro-labor policies. The enduring
source of increasing incomes for employees is a growing and more efficient economy. As workers become richer so they seek greater participation in the workplace and here is the role for unionism. The goal of the regulatory framework is to facilitate this participation without granting a special group the sort of favors that undermine labor productivity.
References


