The consequences of trade union power erosion

Declining union power would not be an overwhelming cause for concern if not for rising wage inequality and the loss of worker voice

Keywords: union density/coverage, bargaining structure, coordination, macro/micro performance, redistribution, voice

ELEVATOR PITCH

The micro- and macroeconomic effects of the declining power of trade unions have been hotly debated by economists and policymakers, although the empirical evidence does little to suggest that the impact of union decline on economic aggregates and firm performance is an overwhelming cause for concern. That said, the association of declining union power with rising earnings inequality and the loss of an important source of dialogue between workers and their firms have proven more worrisome if no less contentious. Causality issues dog the former association and while the diminution in representative voice seems indisputable any depiction of the non-union workplace as an authoritarian “bleak house” is more caricature than reality.

KEY FINDINGS

Pros

- Trade unions under certain bargaining structures can have favorable macro consequences by being less aggressive in their wage bargaining.
- Trade unions can have favorable micro outcomes by stimulating worker voice.
- Where there are benefits to a long-term relationship between the employer and the worker, trade unions can facilitate contracting.
- Trade unions have historically reduced wage inequality.

Cons

- Trade union monopoly power is bad, and its exercise may lead to a misallocation of resources.
- The basis of pro-productive union effects is vague while there exist alternative, non-union voice mechanisms.
- Governance procedures are not exclusive to union regimes and by design may lower rent-seeking behavior injurious to firm performance.
- Unions may no longer reduce wage inequality or support redistributive policies.

AUTHOR’S MAIN MESSAGE

To the extent that unions have been found to have negative effects on net, their decline might be deemed no cause for concern. However, even in these circumstances, “on net” is not a sufficient guide for policy. Rather than a hands-off approach, the general goal should be to stimulate value-enhancing choices by firms and workers, while limiting the downside of rent-seeking.
MOTIVATION

Union density is in retreat. Data for 24 advanced countries indicate that union density has fallen in 21 out of 24 countries over the last 20 years, and in 22 out of 24 countries in the last 30 years. Even if it is not yet possible to speak of convergence—the Ghent countries (Sweden, Finland, Denmark, and Belgium) in which unions provide or administer unemployment insurance being the main outliers—there has been an unambiguous decline in unionism (Figure 1). Sustained decline can be equated with a diminution in union power, despite pockets of union strength.

This article discusses the consequences of this erosion along macroeconomic and microeconomic contours. Although the evidence on union effects is mixed, it can be argued that union decline may give little immediate cause for concern. Even so, two indicators typically associated with union decline—heightened earnings inequality and a potential shortfall in employee voice—occasion more concern.

DISCUSSION OF PROS AND CONS

Collective bargaining and macroeconomic performance

In discussing the macroeconomic effects of unions, it has been conventional to draw a distinction between union membership, union coverage, and bargaining structure. Union membership means employees are union members, whereas coverage means that their workplace unit has collective bargaining coverage regardless of them being a member or not. Union membership density is the ratio of the number of employees who are members of trade unions to all employees in the population. The union coverage rate refers to the proportion of employees whose terms and conditions at work are determined by
collective rather than individual bargaining. The latter proportion generally exceeds the former because wages negotiated by unions are often applied to non-union workers via extension agreements, although it, too, is in decline. For its part, bargaining structure refers to the level at which wages are determined. It ranges from decentralized bargaining at firm level, through intermediate bargaining arrangements (agreements between industry-wide unions and employers’ associations that establish a floor of wages at the industry level), to centralized bargaining procedures (negotiations between labor and employer confederations that set national wage norms). The three “systems” may be said to apply in Anglo-Saxon, continental European, and Nordic nations, respectively, although membership and typology are in reality more fluid than this. Moreover, a given structure can mask differences in the practice of collective bargaining, such as the degree to which there is coordination in bargaining.

From the outset, union membership density and the union coverage rate were associated with adverse outcomes in contrast with initially more favorable results for bargaining structure and coordination. Focusing on the latter, one important study found evidence of a non-linear relation between bargaining level and the change in employment/unemployment, as well as the Okun Index (the inflation rate plus the unemployment rate), when comparing the period 1965–1973 with 1974–1985 [1]. Other studies, however, have reported that countries with coordinated bargaining structure experienced relatively lower equilibrium unemployment rates, although typically the fitted relation was now linear (rather than hump-shaped).

A more recent review of the coordination literature, embracing the various elements of bargaining structure, examines 28 studies, which it breaks down into 174 sub-studies (where the unit of analysis is the relationship between a specific measure of bargaining coordination and an individual performance measure) [2]. On a simple head count, 45% of the sub-studies support the view that coordination works—either by lowering price inflation, unemployment (or a conflation of the two), or by raising employment and productivity, among other things. But the results vary considerably by outcome indicator. Critically, the more sophisticated the estimation technique employed in the study, the more elusive the empirical relationship between bargaining coordination and economic performance.

Another result is that coordination benefits, where observed, were more likely in the 1970s and 1980s than the 1990s. Further, while initially it was thought that coordinated systems were better able to react to or otherwise absorb shocks, more recent research discounts this purported dynamic benefit, although bargaining coordination may well mitigate the harmful effect of union density on unemployment.

On balance, then, union density and union coverage are associated with unfavorable outcomes, while coordination points more to a reduction in the disadvantages of (strong) unionism than indicating a direct effect on the economic aggregates.

All this is rather thin gruel. But an interesting development—the contingency hypothesis—argues that the success of coordination (and centralization for that matter) is contingent on the governance capacity of the bargaining parties at higher levels to bind lower levels (so-called vertical coordination) [3]. Compliance of lower-level actors is facilitated by state provisions for the legal enforceability of collective agreements and a peace obligation during the validity of a collective agreement. Centralized and/or coordinated wage
bargaining, so the argument runs, can only be expected to deliver the macroeconomic goods in conjunction with a high degree of bargaining governability. There is some cross-section empirical evidence favoring this contingency hypothesis in terms of lower inflation and labor costs.

An extension of the contingency argument allows for a more comprehensive categorization of bargaining coordination and governability, or lack thereof, by considering the type of collective agreement for each level of bargaining obtaining at firm level—as well as individual bargaining. This hybridization model has suggested that coordinated sector collective bargaining, governed company and sectoral bargaining, and governed national, sectoral, and company-level agreements (identified with Austria, Germany, and the Nordic countries, respectively) are associated with superior relative labor productivity whereas company and individual bargaining regimes post only an average performance rating compared with the other categories.

But there are problems with both approaches. As far as the baseline contingency model is concerned it is not clear that governance capacity is the most important enabling factor at work here as opposed to, say, the stance of monetary policy. For its part, the hybrid collective bargaining model seemingly does not prove enlightening in understanding other outcomes of a behavioral nature such as the industrial relations climate, employee motivation, or strike incidence.

Decentralization of collective bargaining has been a hallmark of industrial relations since the 1980s. This phenomenon can be considered as equal in importance to deunionization. It includes the tendency for national (i.e. cross industry/inter-sectoral) bargaining to give way to sectoral bargaining as well as an increase in importance of local or enterprise bargaining either at the expense of sectoral bargaining or by acting as an additional layer of bargaining. Expressed another way, coordination and decentralization are the key issues for modern research into the design of fit-for-purpose collective bargaining institutions from the perspective of macroeconomic (and microeconomic) flexibility. Modern theoretical developments in economics have generally applauded the decentralization of collective bargaining—even if industrial relations scholars have been preoccupied with a distinction between organized and disorganized decentralization—while at the same time recognizing that sectoral bargaining with opt-outs and rule setting (rather than uniform wage changes) under multiemployer agreements can mimic the results of decentralized bargaining models. A more integrationist or ecumenical point would be that researchers should have an eye to the design features of bargaining systems that are more or less helpful in achieving flexibility. But the fact remains that this restructuring has, with certain notable exceptions [4], been accorded insufficient attention in the empirical literature, as has the uncertainty introduced by increasing globalization.

Collective bargaining and microeconomic performance

From the perspective of micro theory, union decline again poses a mix of positive and negative elements. The conventional monopoly theory of unions sees their effects as unabashedly negative. Viewed as combinations in restraint of trade, unions introduce distortions into what would otherwise be efficient labor markets. They distort labor market outcomes owing to the increase in compensation above competitive levels. Deadweight losses arise, with too little output being produced in the union sector and
too much in the non-union sector. To these losses in welfare, it is conventional to add the output costs stemming from the union rule-book and reduced management discretion, although these are seldom quantified. Note also that strike costs are little mentioned today because of substitution possibilities, both inter-firm and inter-temporal, and the likelihood that strike threat power will be manifested in the wage premium rather than work stoppages.

But there is a countervailing face of unions that emphasizes their value-enhancing effects. The chief exponents of this collective voice view of unionism note the ambiguity introduced by long-term attachments between the firm and much of its labor force for the efficiency properties of the standard quit or exit mechanism [5]. Reliance by the firm on quits or exit interviews to extract information relevant to the design of an efficient mix of wages and working conditions may introduce inefficiencies by focusing on the preferences of the marginal worker rather than those of more stable and potentially more valuable employees. Collective voice may then outperform individual voice as a means of bringing actual and desired conditions closer together. Crucial to this argument is that many working conditions are public goods, with the implication that they will be underprovided without some form of collective agency—at all times equated in this model with autonomous unions. The same public goods argument can be applied to the supply of effort, assuming there are significant complementarities in worker effort inputs. Based on these public goods arguments, collective voice may therefore lower quits and increase output.

There is also the issue of governance, which refers to the policing and/or monitoring of incomplete employment contracts. Here the collective voice model is consistent with modern contract theory. Assuming that unions make it easier (less costly) to negotiate and administer a governance apparatus, they may be expected to facilitate long-term efficient contracting in a number of ways. For example, a union specializing in information about the contract and in the representation of workers can prevent employers from behaving opportunistically where the reputation effects mechanism—punishing firms that renege on their ex ante promises to take workers’ interests into account by having to pay permanently higher wages—is weak, or indeed where those promises are simply not credible. As a result, employment levels might be higher under unionism. One fly in the ointment, however, is that the union governance argument also hinges on bargaining power and with it a union holdup problem. Another is of course that the reputation effects mechanism might be alive and well.

Empirical evidence for the US, surveyed in an influential review, does not however encourage a sanguine view of this modern perspective on unionism [6].

First, as far as the keynote productivity variable is concerned, union effects are close to zero on average, and at most modestly positive. Second, unions have little direct effect on productivity growth; the lower growth of union firms, after controlling for union–non-union differences in capital and other factors of production, is the consequence of their being located in slower-growing sectors (but see below).

Third, the findings with respect to profitability are of concern. In one sense, a negative profitability effect is to be expected, given a substantial union wage premium in conjunction with almost no effect on productivity. And virtually all US studies point to lower profitability in union regimes, irrespective of the profit measure used. At issue,
however, is the source of the union gain. If the process is merely due to redistribution, there are no implications for efficiency. But there is little to suggest that concentration-related profits are an important source of the gain. More potent sources are current earnings associated with limited foreign competition and growing firm/industry demand.

Fourth, even greater concern is occasioned by union effects on investments in tangible (i.e. investment) and intangible (i.e. research and development, or R&D) capital. US research indicates that unions capture some share of the quasi rents that make up the normal returns on investment in long-lived capital and R&D. Firms rationally seek to limit their exposure to this holdup problem, most obviously by cutting back on these investments. There are both direct and indirect union effects: the former are caused by the union wage tax, while the latter stem from the reduction in profits (relevant because of imperfect capital markets).

Finally, lower profits and investment are manifested in lower employment growth, although infrequently in higher failure rates.

In a rare departure from these pessimistic findings, one US study examining the effects on labor productivity of various working practices, information technology, and management procedures in conjunction with unionism offers a brighter scenario [7]. Specifically, it reports that a hypothetical union plant embracing benchmarking and total quality management, with 50% of its workers meeting on a regular basis (a measure of employee involvement), and operating profit-sharing for its non-managerial employees, would have 13.5% higher labor productivity than a non-union plant with none of these practices. By contrast, the corresponding differential for a high-performance non-union plant is put at only 4.5%. An important qualification, however, is the word “hypothetical,” since such innovative union plants constitute a tiny share of union workplaces in the study sample.

Problems also attach to an innovative US study of the effect of (new) unionization on productivity, output, business survival, and wages [8]. Using a regression discontinuity research design that compares outcomes for employers where unions barely won representation election selections with those where they barely lost, the study reports a negligible impact of unionism on all four outcomes. However, data issues and technical problems stemming from significant discontinuities in the underlying characteristics of establishments at the threshold call into question the validity of the test procedure which in this application is arguably not informative of the average treatment effect of new unions.

But to what extent do these largely negative US results carry over to other countries? After all, most studies confirm that the US union premium for the private sector is unusually high compared with that in other nations.

Cross-country surveys do in fact often report different results for other countries. In particular, the innovation results and (to a lesser extent) the profit results are generally not found for other nations. Given that the data in these studies are rather dated, however, this section concludes with some updated results for Britain and Germany. The British case is interesting because of the shift in the impact of British unions in the 1990s and beyond compared with the 1980s. The evidence points to a reduction in the disadvantages of unionism rather than a reversal for most outcomes—the main exception being profitability where there is the suggestion of a straight reversal of past negative effects.
After US and British research, union and worker representation effects on performance have perhaps been most studied for Germany. Research has focused more on works council than union effects, although recently the two have been examined together (appropriately so, given the dual system of industrial relations in Germany, with collective bargaining typically being conducted at industry level and worker representation at plant level through the agency of works councils). German works councils are the exemplars of collective voice, given their statutory rights (to information, consultation, and codetermination) and constraints (they cannot bargain about terms usually fixed under collective agreements at industry level, and they cannot engage in strike action). But the breadth of their authority inevitably conveys power, and how this is exercised will determine their effects on performance.

Recent German studies exploiting large nationally representative data sets offer some indication that the effect of works councils on firm productivity and even innovation may be positive if the entity is firmly embedded in the dual system (i.e. covered by a sectoral agreement). However, with the pronounced decline in unionism, German sectoral collective bargaining has significantly decentralized, and works councils have come to enjoy formal bargaining rights. One form of this decentralization has been the growth in pacts for employment and competitiveness. These pacts are an outgrowth of opening clauses that have allowed firms to deviate from the normatively binding terms of collective agreements. They are no longer limited to companies in crisis and may be now described as a normal regulatory instrument at a time when collective bargaining standards are guidelines that give firms discretion to arrive at company-specific solutions. What has been the effect of these instruments? The jury is still out, but there are again some grounds for optimism. For example, a recent study has examined the impact of pacts on six firm outcomes—wages, employment, investment, productivity, innovation, and survivability—where the controls are establishments that negotiated over pacts but failed to reach an agreement on their implementation [9]. Despite the diversity in firm behavior and short time period examined, there is no suggestion that pacts negatively impacted any of the selected outcome indicators. Indeed, some robust improvements in employment and innovation are recorded alongside increases in average wages.

The earnings distribution and worker voice

Their redistributive function has sometimes led to unions being described as a sword of justice. Also, apart from the issue of industrial democracy, workers possess valuable private information that is more likely to be disclosed under collective action. Might not the decline in unionism therefore have worrisome implications for inequality and worker voice?

The erosion of union power charted in Figure 1 has been accompanied by a pronounced increase in inequality. It has been well documented in the US and other Anglo-Saxon countries that unions are associated with reductions in earnings inequality. It is but a short step to argue that reductions in the union premium, fewer workers in the union sector, and a diminution of spillover benefits (including an erosion of minimum wages) from the union sector to the (growing) non-union sector—each the consequence of declining union density—underpin the observed increase in inequality. Indeed, a well-known US study has claimed that such factors can explain one-third of the rise in wage inequality among men and one-fifth of the rise in inequality among women from 1972 to 2007 [10].
However, this has not been the dominant narrative among economists. Rather, skill-biased technological change (SBTC) is commonly regarded as the fundamental cause of the rise in income inequality. That is, the nature of scientific progress is said to have strengthened the hand of the most skilled workers (who are depicted as exiting unions) while concurrently the demand for other workers has stagnated. However, a fuller representation of the argument since the 1980s would recognize that inequality has risen due to the hollowing out (that is a reduction in the employment share) of the middle of the jobs distribution in both the US and the EU as a whole.

Further, a recent US study using micro-level data that permits the authors to study union effects over a much longer period than existing work, has qualified the baseline variant of orthodoxy, noting among other things that union density has throughout the sample period 1936–2016 been inversely correlated with the relative skill of union members [11]. That is, union members are relatively more highly skilled today than in the heyday years of the 1950s and 1960s when unions were at their strongest yet made up of relatively less skilled workers.

Nevertheless, although unions may have played a material role in narrowing inequality at their peak or at times of strength both within and beyond the ranks of their members, any such role since the 1980s and 1990s seems unlikely against the more recent backdrop of job polarization. Also, the fact noted earlier that union members are relatively more highly skilled today than in their heyday years suggests diminished support for redistributive policy interventions on the part of organized labor—for which argument there is some support from a recent comparative study of OECD nations [12].

Policy prescriptions hinge on identifying the contributions of market forces and institutional erosion to widening inequality. If the former were to prove predominant after all, then the emphasis should be upon post-market policies, whereas if the latter are more influential a case can be made for measures that include facilitating union formation. Difficult choices are further complicated by the existence of firm product market power and the fact that income distribution matters for growth. In the latter context, however, the received wisdom that greater equality would favor economic growth seems undercut by the recent finding of the reverse association, at least for richer nations.

In light of the above, one is perhaps on firmer ground in speaking of a shortfall of worker voice in the wake of union decline. This topic has generated much debate in the US because of (i) its vanguard position in that retreat, (ii) the pedigree of the collective voice model, and (iii) the results of two large-scale surveys (for 1977 and 1995) indicating that workers desire more voice and influence in the workplace. Thus, around one-third of the non-union, non-managerial workforce in both surveys claimed they would vote for a union if presented with the opportunity, while in 1995 some 27.5% (53%) of all workers reported a sizable (or some) deficit between their actual say on eight workplace issues and what they deemed to be an appropriate say.

An updated and wider ranging survey for 2017 gives comparable estimates of the two shortfalls of 21.4% and 46.9%, respectively, while indicating that rather more non-union workers than before—now 49%—would vote for a union [13]. However, in examining worker satisfaction with independent and internal voice options—where the former include unions, occupational associations, and joining strike action, and the latter encompass conversations with a supervisor, filing a grievance at the workplace, and
joining an employee-manager committee—the new survey also reveals that no “one-sized shoe” fits all workers. That is, some workers are more likely to favor internal options than unions, or to regard each option as superior for some issues but not others.

Other research using the European Company Survey offers support for the US findings in also reporting material representation gaps at the workplace and in identifying alternative options to union voice in the 28 nations of the EU [14].

Research for Britain confirms the US result that the decline in union voice has been accompanied by a significant expansion in non-union voice. Indeed, in the British case, the overall coverage of voice mechanisms has remained high and stable [15]. In short, British employers have thus chosen non-union voice rather than opt for no voice at all.

At issue, then, is the precise size of the gap and the measures required to close it. Although it does not follow that autonomous unionism is the remedy, either from a worker perspective (where workers seek non-adversarial representation) or the practical interest of fostering higher productivity, it has a role to play. Employer-created non-union forms offer the prospects of meeting the aspirations of workers, and of yielding gains to workers and firms alike, but they should not preclude experimentation with conditional deregulation or shifting to a new organizing model.

LIMITATIONS AND GAPS

Despite the wealth of evidence reviewed here, it pertains only to OECD countries and, within that firmament, the majority of studies cover Anglo-Saxon countries, whose experiences may differ in important respects from the rest. The representativeness of the results is therefore in question. Also, understanding of the relationships uncovered in the existing sample of countries is often fragile. Examples are the difficulty of measuring levels of and changes in bargaining structure, and the ambiguity surrounding transforming industrial relations practices and firm performance. Issues of causality loom particularly large. Many of the relationships examined here are descriptive and have a basis in cross-section analysis. There is a pressing need for use of better (high-dimensional) data at the micro level allowing controls for firm and worker fixed effects, in the absence of which spurious attributions of the direction of causality are all too easily made.

SUMMARY AND POLICY ADVICE

Unions have both beneficial and harmful effects in theory.

Union density/coverage is associated with adverse macro outcomes, and bargaining structure/coordination no longer appears to have a direct effect on performance (although it may moderate the harmful effects associated with the former indicators). The potential of bargaining discipline awaits formal validation.

At the micro level, apart from the frankly negative findings for the US, findings for other countries are more nuanced. For its part, the British evidence points to a decline in the disadvantages of unionism rather than a simple reversal of unionism’s negative effects. The unusual notion, encountered in US and British literatures, that adopting transforming industrial relations practices enables union firms to outcompete non-union firms with the same set of practices requires validation. On the other hand, while sectoral bargaining in
Germany has often been criticized on macro grounds, some recent evidence suggests that the country’s dual system of industrial relations is associated with improved productivity and other outcomes.

There is emerging concern over the distributional consequences of union decline. However, there is disagreement over the causes of heightened income inequality, and most economists continue to view SBTC as the culprit. Policy prescriptions involving market-oriented policies are hampered by this division. Somewhat less controversial is the potential shortfall in employee voice attendant upon union decline, which is a real cause for concern in the US because of that nation’s labor laws. And although some encouraging signs of a growth in non-union voice are evident in the data, union voice can undoubtedly help in addressing the challenges posed by a changing world of work, among other things.

The goal of policy should be to stimulate value-enhancing choices by firms and workers while limiting rent-seeking. The emphasis should be upon experimentation and self-regulation, whereby a variety of systems, including the union option, are put up for adoption by the market.

Acknowledgments

The author thanks an anonymous referee and the IZA World of Labor editors for their helpful comments on an earlier draft. The author also thanks Barry Hirsch, Stanley Siebert, and Foteini Tzachrista. Version 2 of the article revises the discussion on inequality and employee voice, updates the figures, and adds new “Key references” [4], [8], [9], [10], [11], [12], [13], [14], [15].

Competing interests

The IZA World of Labor project is committed to the IZA Code of Conduct. The author declares to have observed the principles outlined in the code.

© John T. Addison
REFERENCES

Further reading

Key references

Online extras
The full reference list for this article is available from:
https://wol.iza.org/articles/consequences-of-trade-union-power-erosion
View the evidence map for this article:
https://wol.iza.org/articles/consequences-of-trade-union-power-erosion/map