1 Introduction

John T. Addison and Claus Schnabel

This Handbook seeks to give the reader a sense of the drama in the everyday workings of modern labour markets. There is no better illustration of this than the situation confronted by the trade union. Unions have been buffeted by competitive pressures more than most other institutions. At one level – where unions are combinations in restraint of trade – this result is neither surprising nor unwelcome. But unions are more than this. The twin aims of this volume are to demonstrate the various facets of unionism and thence to address the consequences of what it is that unions do. It is then left up to the reader to form a judgement on the institution and to assess to the seriousness of union decline (or indeed otherwise) in his or her country. Necessarily, the contributors to this volume do not share a common view of the union institution or indeed its fate.

Although trade unions have long been important economic and political actors in industrialized countries, academic interest in them has been thin historically. This was especially true of the economics discipline. As late as 1975, in his short survey of the economic analysis of the union institution, Johnson (1975, p. 23) stated that ‘[t]he study of the behavior and effects of trade unions is not currently one of the major growth industries of the economics profession’. And yet, after surveying the union literature in the mid-1980s, Oswald (1985), Farber (1986) and Hirsch and Addison (1986) all stressed the remarkable growth of work in the area and the substantial progress that had been made in the analysis of unions and union behaviour (see also Freeman and Medoff, 1984). Subsequent advances in union research were identified by Pencavel (1991) and Booth (1995), but such has been the pace and breadth of research in recent years that presentation of a comprehensive, unified description of the current state of play is beyond the capabilities of a single author.

In the light of this embarrassment of riches, we here follow a different strategy, having commissioned chapters from the leading specialists in the field. This Handbook approach has the major advantage that the main issues can be covered in detail and within a framework that tackles different country settings and recognizes contributions from several disciplines. And although this volume will be consulted mainly by students of labour economics, industrial relations, and political science, most of the chapters are written in a rather non-technical way which should also allow non-specialists more than a
glimpse of the current state of research on trade unions. In sum, the Handbook provides a detailed review and analysis of what the editors believe to be the main areas of research in the (economic) analysis of unions, ranging from the determinants and the extent of unions through analysis of their micro- and macro-economic effects to discussion of some key institutional developments at country and regional levels.

Because the existence and the power of trade unions depends on their ability to attract and nurture a loyal membership, Claus Schnabel leads off in Chapter 2 with a survey of the determinants of union membership. The starting point is the conventional demand and supply framework used by economists to analyse the forces that influence union membership, extended to cover the free-rider problem. The key problem for social scientists is to explain why any individual would join a union when membership is costly and where the benefits apply to all workers regardless of their union status. Recent research points to the existence of social customs and of a minimum critical mass of membership or density below which union existence is not viable, implying that a reduction in union membership caused by temporary shocks can be persistent. In short, the fall in unionization observed in many countries – and documented in Chapter 11 by Jelle Visser – may be almost irreversible. Supplementary explanations of union membership from other disciplines of social sciences are also discussed in this chapter, some of which can be incorporated in the economist’s supply–demand and cost–benefit framework whereas others are more difficult to operationalize.

Schnabel’s review of the international empirical evidence on the determinants of union membership shows that while there are substantial differences in the design and the results of the empirical studies, some consistent patterns do emerge. Concerning macro-determinants, for example, there is some evidence across countries that business cycle factors as well as structural developments (such as changes in the composition of the workforce) play a significant role in explaining short-run changes and long-run trends in union membership. Similarly, individual-level cross-sectional studies have identified a number of micro-determinants (such as personal, occupational and firm characteristics, earnings, attitudes, and social variables) that are associated with the unionization decision. And for their part, cross-national analyses point to the importance of institutional determinants of unionism such as the existence of union-administered unemployment insurance schemes. What is often missing, however, are attempts to integrate the macro- and micro-level findings, and it has proved difficult to build a bridge between the variety of theoretical approaches and the empirical literature on the determinants of unionization.

An important insight of recent theoretical research is that union member-
ship and wages are determined simultaneously, and this is also stressed by Robin Naylor in his chapter surveying economic models of union behaviour. In a novel approach, he adapts the structure–conduct–performance framework from industrial economics to organize a review of the theory of trade unions. Focusing first on conduct (that is, the nature of union objectives and the scope of bargaining), he discusses alternative union utility functions and bargaining models. The right-to-manage outcome is shown to generate higher profits than either efficient or sequential bargaining (for a given level of union influence over the wage), so that firms will prefer to keep employment off the bargaining agenda. Naylor then turns to the issue of bargaining structure and to the literature on unionized oligopoly. He shows that in a unionized oligopoly setting there is a dominant strategy towards efficient bargaining. Furthermore, factors such as the product market structure and the nature of the external bargaining environment are important. Product market characteristics—such as the number of firms in the market—affect the outcome of wage bargaining, and Naylor demonstrates that traditional results in oligopoly theory are not always robust to the introduction of union–firm bargaining.

The bargaining process between unions and firms and the emergence of strikes are discussed in Chapter 4 by Peter Cramton and Joseph Tracy. Using non-cooperative bargaining theory and stressing the role of (the firm’s) private information, they present a theoretical model that includes the union’s threat choice between strikes and holdouts (that is extending an expired contract). Since negotiators have an incentive to misrepresent their private information during the bargaining process, labour disputes are interpreted as a (costly) means of communicating this private information. The model provides several insights into how the structure of the labour agreement and the prevailing economic conditions affect the bargaining outcome; for instance, strikes will be more likely after a period of uncompensated inflation and when conditions in the local labour market are tight. Furthermore, in making its threat choice between a strike and a holdout, the union must take into consideration the probability that firms will deploy replacements for striking workers. The model is shown to be consistent with empirical evidence suggesting that strike duration is countercyclical while strike incidence may be pro-cyclical. Cramton and Tracy use the model and its extensions to analyse several labour policies and generate predictions as to their impact on dispute incidence, dispute duration, and the settlement wage. They compare these predictions with extant empirical evidence and evaluate the efficiency and equity ramifications of labour policy. Among other things, they report that cooling-off periods do not seem to affect strike activity whereas bans on replacement workers increase strike frequency and duration as well as wages.
The pay-raising effect of union bargaining demonstrated throughout the Handbook has to be funded from somewhere – from greater productivity, or lower profits, or higher prices. In Chapter 5, David Metcalf examines in detail the effects of unions on productivity, profitability, and investment. He shows that it is not possible to use theory to predict unambiguously the direction of the union effect on productivity. This is because unions can either enhance productivity (by reason of their monitoring role or collective voice effects) or detract from it (via restrictive work practices and adversarial industrial relations). The empirical evidence from the six countries analysed is also mixed, confirming that both effects play a role, with different weights at different times and in different institutional settings. In contrast, the evidence on unions’ effects on profits is more clear-cut, indicating that profits or financial performance are inferior in unionized workplaces, firms, and sectors. For both productivity and profitability, multi-unionism seems to deliver the worst results.

Concerning investment, rent-seeking behaviour by unions might suggest that their impact is negative, but by the same token, firms may also substitute away from expensive labour and invest more in physical capital. The empirical evidence is mixed. It does point to some clear instances of investment-reducing effects of unionism, most notably in North America. But there are also signs that union presence can boost investment in human capital, possibly offsetting the adverse impacts on investment in physical capital where observed. Metcalf stresses the important role of product market competition in both mediating and moderating any adverse union effects, and he emphasizes that cooperation between unions and management yields superior outcomes than adversarial industrial relations.

The macroeconomic effects of unions have occasioned no less controversy than their effects on firm performance. Not surprisingly the Handbook is peppered with references to the debate, and it is the direct focus of Chapter 6. Here Robert Flanagan examines the theory and evidence on the relationship between collective bargaining and macroeconomic performance in industrialized countries. The outcome indicators examined are the usual suspects: inflation and unemployment, a conflation of the two, wage inertia, and the misery index, amongst others. The collective bargaining variable ranges from union density/coverage through degree of centralization to extent of coordination. One important finding to emerge from Flanagan’s review of the basic literature is the sensitivity of the unionism coefficient to sample period and to small changes in the institutional characterization of the country sample. Another is a tendency toward a weakening in beneficial effects of the centralization/coordination covariates in repeated cross section. The author’s response to both tendencies is to look for factors that may underpin these observed changes in the relationships. Thus, for example,
globalization may have removed the hump in the once nonlinear association between degree of centralization in collective bargaining and changes in inflation and unemployment. More fundamentally, however, drawing on the corporatist literature, Flanagan is concerned to show that the effect of an institution on outcomes is likely to be a function of the interactions between collective bargaining and specific policy actions on the part of the authorities and not just the political complexion of governments. In short, he attempts to go behind the stylized representations of the economy in which the standard models are embedded. In the same spirit, he offers an innovative discussion of how European Monetary Union may be expected to affect wage determination under unions.

In the final part of his analysis, Flanagan returns to the empirical regularity that more coordinated bargaining yields improved economic performance to ask whether encouraging greater coordination through institutional reform might be expected to deliver the macro goods. He cautions against any simplistic rejigging of institutions along these lines without greater understanding of the reasons leading contemporary economic actors to seek a greater measure of decentralization in pay bargaining. In his separate discussion of incomes policy, he notes that any such decentralization nevertheless poses an obvious challenge to an effective and necessarily evolving form of incomes policy.

In Chapter 7, David Blanchflower and Alex Bryson cover one of the most familiar areas in the economic analysis of unions, namely, their impact on relative wages. But if the topic is familiar, the authors’ estimates of the course of the union wage premium – or wage gap – are brand new and fill an important gap in the empirical literature. For the USA and the UK, they offer consistent estimates of the overall wage gap through time, using the same format and procedures throughout. Estimates of the union premium by gender, age, race, broad occupational category and sector, and educational attainment are also provided for both countries; and time trends in these differentials are investigated. Next, the cyclical nature of the wage gap is examined for both countries. Finally, the authors provide estimates of the union premium in 17 other countries again using micro datasets.

What are Blanchflower and Bryson’s main findings? To whet the appetite, here are just five. First, the US union premium has held up well – averaging 18 per cent – despite falling union density and coverage. The premium seems to have declined only after 1995, when the economy was booming. Second, to quote the authors, the magnitude of the US union wage differential is today a ‘major liability to the future development of unionism’ in that country. Third, depending on the micro data set used, the union premium in Britain averaged either 10 or 14 per cent over the sample period, but it has collapsed since the mid-1990s such that for a number of
groups – for example, men and private-sector employees – the union coefficient in a log wage equation is no longer statistically significant. Fourth, for both countries the wage gap is countercyclical: as unemployment increases, the union wage premium rises. This leads the authors to speculate that the recent decline in the premium is temporary rather than a trend change in union impact. Finally, the union differential estimated for the other (17) countries examined shows that their unions, too, are able to raise wages well above the equivalent non-union wage.

Chapter 8 is given over to a detailed investigation of the effects of unions on the wage structure. Here David Card, Thomas Lemieux, and Craig Riddell provide an extensive review of the literature before offering a re-analysis of their own. The starting point is methodological, namely, a formal characterization of the potential effects of unions on wage inequality. Using a simple homogeneous worker, two-sector model, the effect on unions on the variance of wages is broken down into a ‘within-sector’ effect (wage dispersion is different in the union and non-union sectors) and a ‘between-sector’ effect (the wedge between union and non-union wages) which is always disequalizing. The model is then extended in two ways. The first complication is to introduce different skill groups. The second extension is to allow for unobserved worker heterogeneity, in recognition that selection might underpin any observed flattening of the wage structure under unionism.

If anything the early literature pointed to a disequalizing effect of unionism on the wage structure, further contributing to the hostility with which unions were often viewed. After 1975, however, the evidence increasingly pointed in an opposite direction, with negative within-sector effects dominating the positive between-sector effect. This equalizing tendency is also reported in second-generation exercises that take account of differences in union coverage and wage premia across different skill groups. And the same is true of smart studies that seek to correct for unobserved skill differences. Card, Lemieux, and Riddell give chapter and verse on this evolving literature and conclude with some important new empirical findings on the male and female wage distributions using large and comparable micro data sets for Canada, the UK, and the USA. They provide information on the raw and skill-adjusted union wage gap or premium, the standard deviation of log union and non-union wages, and a decomposition of variance (i.e. within-sector and between-sector effects). Among their most important findings are the following. First, unions always reduce wage dispersion among male workers, albeit less so when individual characteristics are controlled for. Second, unions do not reduce wage inequality among female workers, partly because unionized women are clustered in the upper end of the wage distribution, and also because their premium is not decreasing in
skill level and is larger than that of males. Third, for the USA and the UK – although not Canada – the decline in unionization is associated with a \textit{ceteris paribus} increase in wage inequality. According to the authors’ estimates, the decline in the fraction of the workforce organized and in the union premium accounted for 31 (14) per cent of the growth in US wage inequality between 1973/74 and 2001 in the simple two-sector (skill-adjusted) model. For the UK, this time for the period 1983–2002, the corresponding values are 29 and 9 per cent, respectively.

The impact of trade unions on innovation is the topic of Chapter 9 by Naercio Menezes-Filho and John van Reenen. Although unions may affect innovation through their effects on relative factor prices and profitability and their stance on the introduction of new technology, recent theoretical work has focused on the possibility that unions will ‘hold up’ firms by expropriating sunk R&D investments through demanding higher wages. This effect may be mitigated or exacerbated by strategic incentives to compete in R&D races. In an attempt to resolve the theoretical ambiguity, Menezes-Filho and van Reenen survey recent micro-econometric studies in the areas of R&D, technological innovation, diffusion and productivity growth. While North American studies find consistently strong and negative impacts of unions on R&D, European studies generally do not uncover such negative effects. There is no consensus concerning the impact of unions on innovation, diffusion, and productivity growth. The authors conclude that the knowledge base they draw on is much thinner than in other areas of union research, which may be due to unsolved econometric problems or to institutional differences between nations in union attitudes and ability to bargain.

Although several chapters touch upon the unions and politics nexus, the analysis of Wolfgang Streeck and Anke Hassel in Chapter 10 centres on this theme. Thus, the authors first offer a typology of unions and politics along the two dimensions of political unity and politicization – a framework constructed on the basis of the historical experiences of a sample of more than a dozen countries. It is shown that the manner in which unions achieve political influence varies widely. It may occur through political exchange as when centralized unions achieve concessions from governments under incomes policy in return for wage discipline, or through repeated interaction with a given political party, or via functional representation on national economic policy councils and in the joint administration of social security schemes. Finally, there is lobbying, which is of increasing importance as the formal links between unions and (centre-left) political parties have become ever more tenuous.

Just as with the ‘mechanisms’, there have occurred changes through time in the union role in economic policy. Among the more recent was the abrupt
diminution in that role with the dismantling of incomes policy under the challenge of monetarism. (Vulgo: if there was no output–inflation tradeoff to exploit, what was to be gained by negotiating with unions?) But Streeck and Hassel argue that this development was to prove more dramatic than sustained, with the recognition that the effects of monetarist policies were not independent of the structure of wage bargaining. Moreover, other union functions – such as their training capacity and potential for productivity improvement via collective voice – were often to tip the balance in favour of their continued inclusion in economic policy councils (in continental Europe if not the Anglo-Saxon OECD countries). Finally, the authors see unions as fundamental to development of the welfare state, arguing that the process of industrialization and modernization both enabled and required such public welfare provision. Specifically, they refer to a ‘decommodification’ of labour – meaning the extent to which society protects the individual from the market. The degree of decommodification is shown to be greatest where bargaining is centralized.

Yet the current welfare state is in retrenchment, with implications for unionism both in terms of the social wage and joint governance. Although the authors view this development as temporary, they recognize that the viability of the unions as political actors is under long-term pressure. Quite apart from the challenges of globalization, unions are less representative of the working population than heretofore. This calls into question their legitimacy as political representatives of the workforce. Moreover, the core membership is increasingly a narrow constituency. If it is to survive, corporatism has to reflect societal and not just sectional interests.

While most chapters in this Handbook draw from the experience of several countries, Jelle Visser in Chapter 11 ranges much further afield. He covers developments in more than 100 nations. Combining union membership data from various sources, Visser estimates the total number of union members at around 320 million. This translates into a global union density rate of 23 per cent if we exclude the one-half of the world’s labour force that is involved in agricultural activities and usually not represented by trade unions. Visser discusses developments in unionism and unionization across different world regions, in most cases comparing the mid-1980s with the mid-1990s. He shows that union density has fallen substantially not only in the advanced industrialized countries of Western Europe, North America and Oceania, but also in the transition countries of Central and Eastern Europe where the new market economy is mostly non-union. Union retrenchment is also the general trend in Central and Latin America (with the notable exception of Chile after its return to democracy). In Asia the picture is more divergent, ranging from strong increases in union density in Hong Kong to dramatic membership losses in Thailand and
India, but the sheer size of the continent and the profound differences between its various parts and countries defy meaningful generalizations. The same is true for the Middle East and Africa, but despite serious data problems it becomes clear that we observe in Africa the world’s lowest unionization rates. For a smaller group of countries, Visser is able to present data on the composition of unionization by sector, gender, and working hours. Among other things, this material indicates that women and part-time workers are often underrepresented in unions. Finally, he shows that although the percentage of workers covered by collective agreements can exceed the proportion unionized, union coverage no less than union density exhibits decline in many countries.

Returning to advanced industrialized economies, perhaps no other country in recent years has witnessed greater change in its industrial relations framework than the UK. In their wide-ranging chapter, John Addison and Stanley Siebert describe the dramatic developments and their consequences. Like Gaul, their treatment is in three parts. The first part charts the six major pieces of legislation – conventionally described as ‘anti-union’ – that were enacted by successive Conservative administrations between 1980 and 1993, and links them to the subsequent decline in unionism and to changes in firm performance and that of the macro-economy. In addition, the consequences of union decline for the wage structure are also accorded special attention. The second part examines the accession of ‘New Labour’ and reviews its domestic reform agenda, today largely in place. That agenda comprises two general pieces of employment and employment relations law plus a new national minimum wage. At first (and second) blush these changes do not return Britain to the mid-1970s even if they do imply an increase in union membership and rising costs for business, other things equal. For evidence of more profound change one has to turn to the third part of the story: the social policy agenda of the European Union. Almost immediately upon taking office, New Labour signed up to the Social Chapter. This means that a slew of new legislation seeking to regulate the employment relation (mostly decided by qualified majority) is now in immediate prospect. Europe is therefore set to impact the theory and practice of British industrial relations. The authors provide chapter and verse on the actual and prospective legislation.

The Europeanization of British industrial relations is one thing, the emergence of European-wide collective bargaining potentially quite another. The latter issue is the subject of Chapter 13, in which Dieter Sadowski, Oliver Ludewig, and Florian Turk argue that pan-European bargaining is at best a long way off for reasons that reflect the choices of the parties. At one level, Europeanization can be indexed by the social policy initiatives of the European Union (EU), and in particular those stemming
from the uprated ‘social dialogue’ between the two sides of industry at European (and sectoral) level in the wake of the Maastricht Treaty. The authors characterize this process as ‘Commission-driven’. The other type of Europeanization identified is ‘spontaneous’ initiatives: instances of cross-border collective bargaining and networks of the institutions formed by metalworker unions. These developments in turn helped form the peak organizations of labour and capital, or social partners (chief among which are UNICE/CEEP on the employer side and the ETUC on the labour side).

To establish their point that Europeanization is at this stage a fragile creature, Sadowski, Ludewig, and Turk offer the notion of an optimal bargaining area as informed by the economic theory of clubs. The goal is to determine (a) whether the actors have an interest in the Europeanization of collective bargaining, (b) the type of Europeanization that will be preferred, and (c) the topological features of the collective bargaining system that will evolve. There follows a detailed discussion of the positioning of the respective social partners and of government institutions. The conclusion of the chapter is that there is and will be an increasing European dimension of collective bargaining, but that this process does not require a ‘standardization of collective bargaining and the transnational unification of social actors’. Rather, we should expect to see a centralization of information sharing, within-company bargaining in European multinationals, some wage coordination, and instances of cross-border bargaining accentuated by the eastern enlargement of the EU.

In the final chapter of this volume, John Delaney reviews current and likely developments in industrial relations in the United States. The picture that he paints is less than reassuring for the union movement of the world’s largest economy. But could it be otherwise given the magnitude of the union wage premium (see Chapter 7), the seeming adverse effect of US unions on firm performance (see Chapters 5 and 9), and the critical mass problem associated with the long-standing decline in unionism (see Chapter 7 and also Chapters 2 and 3). The figures bear repeating: as of 2001, union density in the USA was just 13.5 per cent – 9 per cent in the private sector and 37.4 per cent in the public sector. Delaney charts the by now familiar problem of unions not being able to organize sufficient new members to offset natural attrition. In doing so, he accepts that there is no great clamour for organization from those workers that have most to gain from unionization. Is the solution for unions increased engagement of the polity? Possibly, we are told, but campaign finance reform may block this avenue. Thus far at least the exertion of political power has been used to stop changes in the law hostile to union interests, but has proved inadequate to change the law in favour of unions. The stillborn recommendations of the Dunlop Commission (1994a, 1994b) are eloquent testimony of this. In
common with other observers, Delaney sees the National Labor Relations Act (NLRA) as something of a problem for the union movement. (The reasons are technical and include ineffective penalties for unfair labour practices on the part of employers, and section 8 (a) 2 of the NLRA which may remove swathes of employees from its protection.) But, reflecting his recognition of the role of market forces, Delaney does not really argue that American labour law has failed. Rather, he argues that the NLRA is becoming increasingly irrelevant to employees with the change in the nature of work and as many jobs gain supervisory and managerial duties. Unions have to address this challenge by organizing outside of the NLRA and by offering a different product. He is evidently thinking of a cooperative unionism, a closer partnership between management and labour, very much on the lines discussed by David Metcalf in Chapter 5.

This, then, is our product. Necessarily, it is incomplete. Space constraints have resulted in the exclusion of some countries and regions that would otherwise undoubtedly merit consideration. The most obvious omission is Japan (see the recent analysis of Tachibanaki and Noda, 2000). Note, however, that data limitations also explain the omission of nations from the country samples examined in this volume. Our response here has been to ask some contributors to state the problem and yet others to provide as much international data as possible on, say, the wage gap or density so as to provide either a resource base or a research stimulus. The effect of incomplete coverage is of course that our generalizations about the behaviour and effect of trade unions must be guarded. And even within the samples investigated here a number of caveats are required. At one level, the progression from theory to evidence is sometimes opaque, meaning that even an agreed set of findings can be consistent with very different positions. At another level, and in addition to standard problems of statistical inference – stemming in particular from the non-random distribution of union presence and lack of information on variables associated with the particular outcome indicator that are also correlated with the union measure – a certain imprecision attaches to some of the union effects charted in this Handbook. We see union effects on innovation as a notable case in point. That said, and as the chapters in this volume will make clear, theoretical work on unions is advancing, the gap between theory and measurement is closing, and the empirics are often state of the art. But note the irony: Johnson (1975), writing during the golden age of unionism, complained about the lack of union studies. Today, at a time when unionization has reached an all-time low, economists and other social scientists can claim a much better understanding of the institution.

Finally, some acknowledgments are in order. First and foremost, we are indebted to the contributors to this volume for their sustained efforts and
for the tonic of some stimulating discussions. We wish also to thank Lena
Koller for her invaluable technical assistance in preparing this Handbook.
Last but not least we are indebted to Luke Adams at Edward Elgar who
encouraged us to pursue this project.

References
Booth, Alison L. (1995), The Economics of the Trade Union, Cambridge: Cambridge
University Press.
Dunlop Commission (1994a), Fact-Finding Report of the Commission on the Future of Worker-
Management Relations, Washington, DC: US Department of Labor, US Department of
Commerce, May.
Dunlop Commission (1994b), Report and Recommendations of the Commission on the Future of
Worker–Management Relations, Washington, DC: US Department of Labor, US
Department of Commerce, December.
Layard (eds), Handbook of Labor Economics, Volume 2, Amsterdam: North Holland, pp.
1039–1089.
Freeman, Richard B. and James L. Medoff (1984), What Do Unions Do?, New York: Basic
Books.
Hirsch, Barry T. and John T. Addison (1986), The Economic Analysis of Unions, Boston, MA.::
Allen and Unwin.
Johnson, George E. (1975), ‘Economic Analysis of Trade Unionism’, American Economic
Tachibanaki, Toshiaki and Tomohiko Noda (2000), The Economic Effects of Trade Unions in