Firm Behavior When Workers Can Unionize

Pablo Ruiz-Verdú*

First draft: March 2002
This version: January 21, 2003

Abstract

Unionization imposes substantial costs on employers. This paper develops a model that recognizes that, as a result, employers will set wages and employment taking into account the effect of their decisions on workers' incentives to organize. This model of employer behavior allows us to address two questions jointly: What determines which firms become unionized? And what are the consequences of unionization for employment and wages in nonunion firms? The implications of the model depart significantly from those of previous work, which either ignored employers' strategic behavior, or treated these questions in isolation.

JEL: J23, J31, J51, J41, M50, D82. Keywords: employment determination, wage determination, unions, wage differentials.

*Address: Universidad Carlos III de Madrid, Department of Business Administration. Calle Madrid, 126 28903 - Getafe, Madrid - Spain. E-mail: prverdu@emp.uc3m.es.

I Introduction

Unionization is costly to employers. Despite the dramatic decline in the rate of union membership in the U.S. over the past twenty years, the wage premium that unions guarantee their members is substantial – between 15 and 25 percent – and has not shown any significant downward trend (see Figure 1[1]). Therefore, unless unionization has a large positive effect on productivity, firms have powerful incentives to remain nonunion. To achieve this goal, they adjust their wage and employment policies to make joining a union less attractive to workers (Rees and Schultz (1970), Foulkes (1980)), and, if that fails, resist union organization strongly, even if this frequently involves taking legally questionable actions (Freeman (1986), Freeman and Kleiner (1990)).[2]

Thus when faced with potential unionization, firms take into account the effect of their decisions on workers' incentives to unionize. This means that, on the one hand, firms' wage and employment choices under the threat of unionization will differ from those they would make in the absence of such threat, and that, on the other hand, unionization cannot be analyzed independently of wage and employment determination. The contribution of this paper is to develop a model in which nonunion firms behave strategically, and in which wages, employment, and the incidence of unionization are simultaneously determined. The model shows that, once we account for this simultaneity, we obtain predictions both about the effects of the threat of collective bargaining on employment and wages, and about the determinants of unionization that depart significantly from the received wisdom.

There are two main views as to the effect of unions on the wages and employment of nonunion workers.[3] According to the first of these views (the *crowding* or *spillover* view), the nonunion sector behaves competitively, and therefore the impact of unionization on the wages and employment of nonunion workers takes place solely through its effect on the supply of labor to the nonunion sector. If, as commonly assumed, collective bargaining reduces employment in the union sector, the effect of unionization will be to lower nonunion wages, as the nonunion sector absorbs the displaced workers. The *union threat* view, in contrast, holds that nonunion firms raise wages to forestall unionization, and, as a consequence cut employment as they move

[1] Figure 1 plots union density and the union wage gap for private sector, nonagricultural workers in the U.S. for the period 1973-1981,1983-1998. The wage differentials are calculated in Hirsch and Schumacher (2000) from CPS data, excluding workers with earnings imputed by the Census. See Hirsch and Schumacher (2000) for a discussion of the dataset, the controls used in the wage equations, and the reasons to exclude allocated earners.

[2] In 1994, the ratio of unfair labor practice complaints issued by the National Labor Relations Board (NLRB) to the number of union recognition elections was 0.843 (Ehrenberg and Smith (1997), p. 493).

[3] See Ehrenberg and Smith (1997) for a standard textbook exposition of these two views.

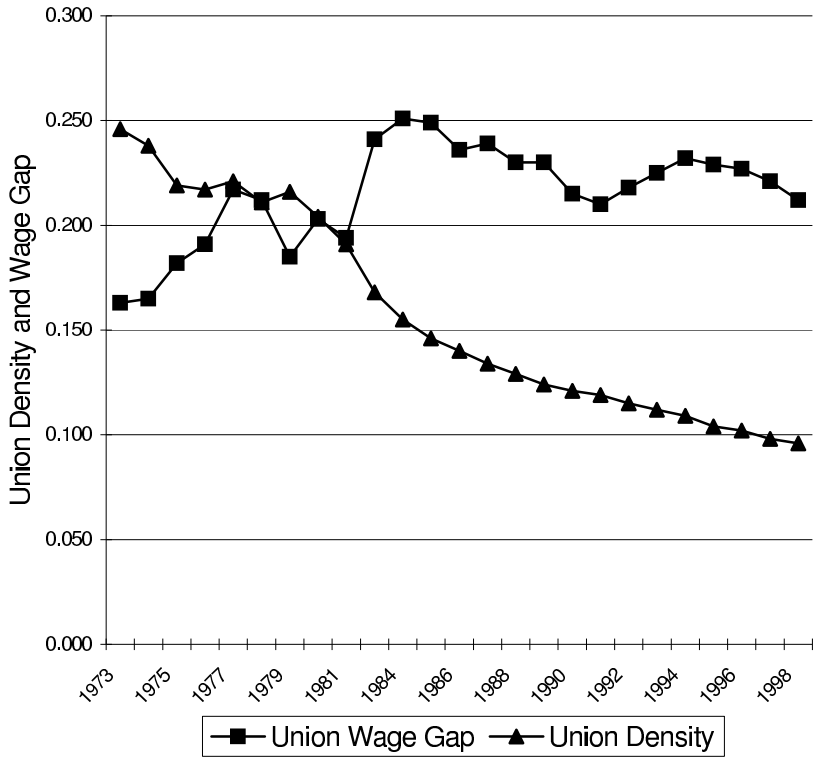

Figure 1: Union-nonunion Wage Gap and Union Density among Private Sector Nonagricultural Workers. 1973-1998. (Source: Hirsch and Schumacher (2000))

up their labor demand curves. Since wages are prevented from falling by the union threat, the labor market does not clear, and unemployment ensues.

With respect to the determinants of the incidence of unionization, it is often argued that those firms where the expected wage gains from unionization are larger are more likely to become unionized. To the extent that changes in estimated union wage premia reflect changes in the expected wage gains from unionization, union membership and the union wage premium should be positively correlated. This argument has led some researchers (e.g. Blanchflower (1997)) to consider it a puzzle that despite the large decline in U.S. union density in the last two decades, union wage premia have remained unaltered. A proposed explanation for this fact (Freeman (1986)) is that the high union wage differentials have, in fact, been the cause of the continuous decline in union density, as they would have led firms to resist organization more strongly. However, this story does not explain what prompted the initial shift to higher wage differentials, and why, after the long decline in union density, wage differentials have remained high rather than return to their initial values.

Unionization is also often regarded as more likely to occur in more profitable firms or in-

dustries, since wage gains are expected to be greater where rents are large (Hirsch (1991)).[4] Whether or not more profitable firms are more likely to become organized is a question of great importance for empirical work, since one of the main problems in the literature examining the effects of unions is that unionization is likely to be correlated with unobservable firm characteristics that also influence the variable of interest (be it profits, wages, etc.). For example, suppose that we were interested in estimating the effect of unionization on profits, and unionization were more likely to occur in more profitable firms. Then, unless we can perfectly control for expected profitability, OLS estimates of the union effect on profits would tend to be downward-biased, that is, they would show a smaller –in absolute value– negative effect. Similarly, estimates of the union wage premium would tend to overestimate the true effect of unions on compensation.

In contrast to these views, the model in this paper shows that: (1) The threat of unionization induces nonunion employers to increase their wages above competitive levels, but does *not* generally reduce employment as a consequence. (2) Only the presence of asymmetric information about firm profitability may lead worse performing firms to reduce employment to convince workers that they cannot expect high wage gains if they organize. (3) Firms earning *lower* rents are more likely to become unionized. (4) Negative industry shocks make unionization more likely. (5) Conventional estimates of the effect of unions on profits will be upward biased, while estimates of the union wage premium may be downward biased. (6) In equilibrium, union membership and the union wage premium will tend to be *negatively* correlated.

These predictions follow from the recognition that firms will behave strategically when faced with the threat of unionization. As long as the process of organization entails some costs to workers or unionization reduces rents, firms will avoid unionization by increasing wages so as to match workers' expected utility from unionization. Therefore, union organization is possible only in those firms where unions improve productive efficiency or provide benefits to workers that employers cannot offer themselves. However, if we allow that employers may be better informed than workers about their firms' prospects, unionization can take place in equilibrium, even if it does not have any positive effect on efficiency. If workers do not know with certainty what they would obtain if they organized, their expectations will be too optimistic in those firms with less than average profitability. These firms may thus find it too costly to prevent unionization, and may allow union formation in order to signal their low profitability.

The increase in wages required to prevent unionization, however, does not reduce employ-

[4]Freeman (1986) and others, however, have questioned this view by pointing at employers' incentives to resist unionization.

ment, as predicted by the standard union threat view. By increasing employment, the firm reduces the wage it needs to pay to dissuade workers from unionizing, because a higher employment level is likely to decrease both the negotiated union wage and workers' probability of employment after organization. This strategic motive to increase employment more than compensates for the rise in wage costs, so that no reduction in employment takes place despite the higher wages. However, if firms have private information about their profitability, less profitable firms may find it optimal to employ fewer workers to distinguish themselves from profitable firms, for which the reduction in employment would be more costly, and thus convince their employees to moderate their wage demands.

While there exists a large empirical literature that attempts to measure the effects of unions on wages and firm performance,[5] the theoretical work on the determinants of unionization and its effects on the behavior of nonunion firms is scarce. Rosen (1969) first introduced a simple reduced-form model of wage determination in which firms take into account the effect of their wage choice on the probability of unionization. Following on Rosen's contribution, Dickens (1986) developed a model of wage and employment determination in the presence of a union threat, and raised the possibility that it may be optimal for nonunion firms to overhire.

With respect to the determinants of unionization, Freeman (1986) and others have tried to explain the extent of unionization as a function of the expected wage gains from collective bargaining, recognizing that these expected wage gains will affect the demand for union services by workers, and employers' incentives to resist unionization in opposite directions.[6] However, these authors take the wage premium as given, and hence cannot account for the effect that the threat of collective action has on the determination of this wage premium. Lazear (1983) provides one of the few models that attempt to simultaneously explain unionization, employment, and wages. In his model, workers announce an industry-wide wage demand, and employers decide whether to accept that demand (thus becoming "unionized") or to resist unionization at a cost. Therefore, employers are given a passive role, while yet-unorganized workers are able to make an industry-wide wage demand. The model thus provides a better description of a unionized industry with centralized bargaining in which employers can deviate from the collective bargaining agreement at a cost, than of an industry with either no or decentralized bargaining, in which workers can organize at the firm level. Finally, a series of papers (Booth (1985), Naylor and Cripps (1993), Corneo (1995)) have analyzed workers' incentives to join unions when the

[5] For surveys of this literature, see Lewis (1986), Pencavel (1991), Hirsch (1991), and Booth (1995).

[6] In a related effort, Farber (1990) attempted to decompose the changes in union density into changes in the demand for and the supply of union services.

benefits of unionization extend to all workers. According to this line of research workers would join unions to benefit from the reputational utility provided by membership.

As mentioned above, when workers can unionize the analysis of unionization and of firms' wage and employment policies cannot proceed separately. In this paper, we develop a model that integrates the insights of the different strands of research just discussed, leading to novel, and in some cases unexpected, predictions. On the one hand, the model analyzes rigorously the impact of workers' ability to unionize on wages and employment, extending the work initiated by Rosen (1969), and, especially Dickens (1986). On the other hand, it shows that taking into account the strategic nature of firm behavior when workers can unionize leads to predictions about the incidence of unionization that differ from commonly held views. Moreover, by studying unionization and wage and employment determination jointly, the model tackles one of the main problems of the literature, both theoretical and empirical, on the determinants and effects of unionization, namely the interdependence of firm policies and unionization decisions.

The paper is organized as follows. Section II presents the model and studies firm behavior and unionization under complete information. Section III, in turn, studies wage and employment determination when firms have private information about their profitability. Section IV discusses the model's implications in light of existing evidence. Finally, section V offers suggestions for future research, and section VI concludes.

II Strategic Wage and Employment Determination and Unionization

A Setup

Consider a firm that lives for two periods. In the first period, the firm offers a wage w, and hires labor, l, from a pool of identical workers, who can earn a wage $\overline{w}$ outside of the firm. During the first period, the hired workers acquire firm-specific skills and produce no output. The firm-specific skills imply that, in the second period, employees are more valuable to the firm than alternative workers. That is, if $R(l)$ and $\overline{R}(l)$ are the revenues the firm can obtain if it employs l skilled and unskilled workers, respectively, then $R(l) > \overline{R}(l)$.[7] R and $\overline{R}$ are increasing, concave and twice continuously differentiable functions, with $R(0) = \overline{R}(0) = 0$. R'

[7] Even in the absence of firm-specific skills, firing and hiring costs may create a wedge between the value to the firm of an existing employee, and that of a new worker. Note that if $R(l) = \overline{R}(l)$, and there were no hiring and firing costs, there would be no room for bargaining.

is also assumed to be concave.[8]

In the second period, once they have acquired specific skills, employees may attempt to renegotiate the wage by unionizing and bargaining collectively with the firm. Unionization, however, is costly for workers, who have to incur organization costs $C(l)$ to be able to form the union and negotiate with the firm. If collective bargaining takes place, a wage $w^u(l)$ is set, and $l^u(l)$ skilled employees remain in the firm, where $l^u(l) \leq l$ since it is not possible to replace workers with specific skills in the second period. The revenues produced by unionized workers, R^u, are allowed to be different from those they would produce absent the union, R.[9] This difference could be negative if the presence of the union led to an inefficient organization of work, say by imposing restrictive work rules, or positive if it provided a communication channel to workers (the so called "voice" effect) or improved morale.[10] At the end of the period, production takes place and wages are paid. Figure 2 represents the model's time-line.

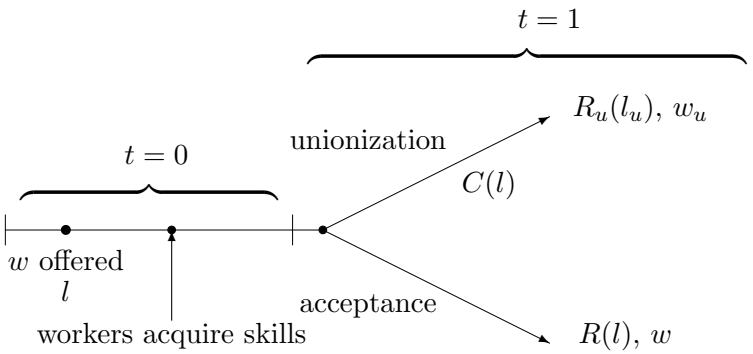

Figure 2: Time-Line

It is worth discussing some of the model's assumptions in greater detail. First, the model assumes that workers start unorganized. This assumption tries to capture how actual labor markets work, at least in the U.S., where, at their inception, firms are generally union-free. Workers' organization decision is more a response to employers' policies than a spontaneous impulse to become union members.

Second, the process of union organization is left unmodeled: workers can organize to bargain collectively simply by incurring organization costs C. A more realistic interpretation of this assumption is that C are the costs workers would incur at the equilibrium of an organization

[8] This assumption is made only for tractability.

[9] R^u is assumed to have the same properties as R.

[10] By giving workers a powerful device to punish deviations from implicit agreements, the strike, the presence of a union could also expand the set of incentive compatible contracts, a change that would affect production efficiency.

game in which they would devote time and effort to organizing, and management would fight the organization attempt. Therefore, C includes not only the direct costs associated with the organizing effort (or the payment to an established union for its services), but also the expected losses due to management's retaliation against union organizers. Modeling this organization game is, however, beyond the scope of this paper.

Third, workers are assumed to be able to extract rents only if they bargain collectively. This is a simplifying assumption: the analysis applies whenever workers' bargaining power is greater if they bargain collectively than if they bargain individually.[11]

Finally, the model assumes that no payments are made and no production takes place in the first period. The implications of this assumption will be discussed after presenting the model's results.

B The Collective Bargaining Stage

If a union is formed, it negotiates with management over wages and, possibly, employment. For simplicity, this paper will assume that the union maximizes the expected utility of its risk neutral workers, and that bargaining takes place according to the monopoly union model. In this model, the union is able to impose its wage demands on employers, who, in turn, keep their prerogative to set employment at will given the wage set by the union. This is perhaps the model of collective bargaining most widely used in applications. Moreover, it has the attractive feature, for our purposes, that collective bargaining outcomes are highly inefficient. The following lemma summarizes its main properties. (All proofs can be found in the appendix.)

Lemma 1 *In the monopoly union model, there exists a level of employment l^{u*} such that:*

1. $l^{u*} < l_c^u = R^{u'-1}(\overline{w})$
2. $l^u(l) = l$, *for* $l \leq l^{u*}$, *and* $l^u(l) = l^{u*}$, *for* $l > l^{u*}$
3. $w^u(l) = R^{u'}(l^u(l)) > \overline{w}$
4. $\Pi_U(l) \equiv R^u(l^u(l)) - l^u(l)w^u(l)$ *is increasing in* l *for* $l \leq l^{u*}$, *and constant for* $l > l^{u*}$.

Therefore, in the monopoly union model, the wage is always on the marginal revenue curve and employment below the competitive level (the level l_c^u that would be chosen if the wage was

[11] See Stole and Zwiebel (1996) and Segal (2001) for a theoretical analysis of individual wage bargaining, and the conditions determining the relative advantage of collective versus individual bargaining.

given by $\overline{w}$). If initial employment is large enough, the union outcome implies $l^u(l) = l^{u*}$; if initial employment is below l^{u*}, the resulting employment level is equal to initial employment. Finally, profits are nondecreasing in initial employment, as profits increase as we move to the right along the marginal revenue curve.

C The Optimality of Union Avoidance

Employers can always guarantee that workers remain nonunion: if they offer a wage greater or equal than workers' expected payoff from unionization, workers would have no incentive to organize. If they offer a lower wage, however, they will trigger unionization. Therefore, unionization will take place only when it is less costly to the firm than union avoidance. To see when this will be the case, it is instructive to abstract from the employment determination problem for a moment. So, let us suppose that employment is fixed and given by l. Then, if workers unionize, profits are $R^u(l) - w^u(l)l$ and workers earn $w^u(l) - \frac{C(l)}{l}$. Therefore, if the firm offers $w = w^u(l) - \frac{C(l)}{l}$, it can avert unionization and earn profits $R(l) - w^u(l)l + C(l)$. It follows that it is optimal for the firm to avoid unionization unless $R^u(l) - C(l) > R(l)$. Put differently, unless the union can increase total surplus, it will be optimal for the firm to offer a wage high enough to convince workers not to unionize.

The following proposition shows that this simple argument also holds if the firm is allowed to determine the level of employment and gives an expression for the wage necessary to prevent union formation.

Proposition 1 *Let l^* be the optimal employment choice, and l^{u*} as defined by Lemma 1. Then:*

i. If

$$R^u(l^{u*}) - C(l^{u*}) \leq R(l^{u*}), \tag{1}$$

at the unique subgame perfect equilibrium, the firm offers $w(l^*) = \max\{\frac{l^u(l^*)}{l^*}w^u(l^*) + \frac{l^* - l^u(l^*)}{l^*}\overline{w} - \frac{C(l^*)}{l^*}, \overline{w}\}$, and workers accept the wage offer.

ii. A necessary but not sufficient condition for unionization is:

$$R^u(l^{u*}) - C(l^{u*}) > R(l^{u*}) \tag{2}$$

Therefore, it is optimal for the firm to avoid unionization as long as it can mimic the best possible union outcome, in the sense that it can guarantee a level of total surplus $(R(l) - \overline{w})$ at

least as high as under the union ($R^u(l^{u*}) - C(l^{u*}) - \overline{w}$). Put differently, *unionization can take place only if it increases total surplus.*

This result, which can be seen as a special case of the Coase Theorem, conflicts with the idea that unions are inefficient either directly, because they harm productivity, (which would imply $R^u < R$) or indirectly, because they distort input choices. According to Proposition 1, the appearance of inefficient unions will be optimally avoided simply by setting high enough wages. In fact, even if unions were merely redistributive, not affecting productivity or input choices, the presence of organization costs would preclude union formation. Allowing unionization is optimal for the firm only if the presence of a union increases rents. This, however, does not imply that firms prefer unionization to the competitive outcome, since, unless the positive effect of unions is substantial, profits will be lower in unionized firms than in nonunion firms paying the alternative wage.

It is worth mentioning that the result that unions can appear only if they increase total surplus is implicitly present in the papers, mentioned in the introduction, that have studied union membership when union wages are a public good. These papers necessitate that the union provide some private utility (reputational utility) to its members beyond the increase in expected income, thus increasing surplus, to obtain union formation in equilibrium.

It should also be noted that, although the organization game is not analyzed in this paper, it would be always better for firms to prevent unionization by increasing the wage than by letting the organization attempt happen and then defeating it. If firms could defeat organization attempts with high probability, the expected payoff from unionization would be low. This would just make it cheap for the employer to prevent unionization by offering a low wage, but the argument behind Proposition 1 would still go through.

D Employment Determination under the Threat of Unionization

Proposition 1 shows that the threat of unionization forces employers to pay workers more than their reservation wage. Because of this increase in the cost of labor, employers will tend to hire less workers, which is the effect usually considered in discussions of the union threat. However, Proposition 1 also shows that the wage required to prevent unionization itself depends on the level of employment. Therefore, employment, and not just the wage, becomes a strategic variable: when workers can unionize, employment will be set taking into account its effect on workers' incentives to join a union. If increasing employment reduces the wage required to prevent unionization, then this strategic motive will counter the employment-reducing effect

of the union threat. Otherwise, employment will be even lower than in standard union threat models.

To simplify the exposition, hereafter unions will be assumed to have no direct effect on productive efficiency, that is, $R^u(l^u) = R(l^u)$, so that the superscript will be omitted. It is important to note that this does not imply that unions cannot be inefficient: they may still affect efficiency greatly by distorting input choices.

Suppose that labor supply is infinitely elastic at wage $\overline{w}$, and let l_c be the level of employment that a firm with revenue function R would choose in the absence of any union threat:

$$l_c = argmax_l\{R(l) - \overline{w}l\}$$

Then, l_c, which will be labelled the competitive outcome, satisfies $R'(l_c) = \overline{w}$.

A firm facing a threat of unionization, however, cannot take the wage as given. Proposition 1 implies that the firm will instead solve the following problem:

$$\begin{aligned} Max_l \quad & R(l) - w(l)l \\ s.t. \quad & w(l) = \max\{\frac{l^u(l)}{l}(w^u(l) - \overline{w}) + \overline{w} - \frac{C(l)}{l}, \overline{w}\}. \end{aligned}$$

Let l^* denote the solution to this problem.

Proposition 2 *If $C(l)$ is nondecreasing in l, then $l^* \geq l_c$ and $w \geq \overline{w}$. If $C(l)$ is strictly increasing in l, and $\frac{l^{u*}}{l_c}(w^{u*} - \overline{w}) + \overline{w} - \frac{C(l_c)}{l_c} > \overline{w}$, then $l^* > l_c$ and $w \geq \overline{w}$.*

Proposition 2 states that *both* the wage and employment are higher under the threat of unionization than in the competitive case, a result that runs counter the two prevailing views of the effect of unions on the wages and employment of nonunion workers discussed in the introduction. On the one hand, once we drop the assumption that employers are wage-takers (and the very presence of unions in an industry should lead us to do so), it becomes difficult for unionization to lower nonunion wages as predicted by the crowding view. To do so, the supply shift due to the workers displaced from the union to the nonunion sector should be large enough to compensate for the upward shift in the wage-employment schedule caused by the union threat. On the other hand, despite raising nonunion wages, the threat of unionization does not reduce employment as predicted by the conventional union threat view. The explanation for this apparently counter-intuitive result is that, because of the threat of unionization, the marginal cost of an extra worker is not equal to the wage, so that employment is not on the

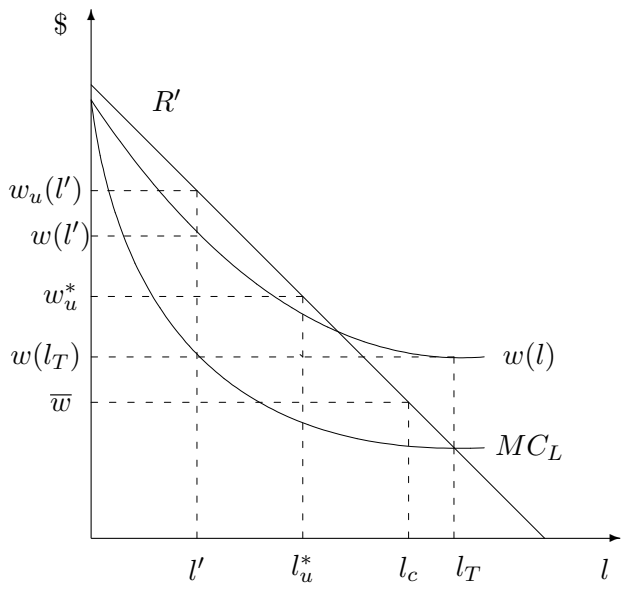

Figure 3: Overemployment in the Monopoly Union Model

competitive labor demand curve. To avoid unionization, employers are forced to set the wage equal to the expected utility from organization. Therefore, if the presence of an extra worker decreases this expected utility, the marginal cost of labor will be less than the going wage. Lemma 1 ensures that this is the case: hiring an extra worker leads either to lower union wages (if $l < l^{u*}$) or to a greater probability of unemployment (if $l > l^{u*}$). In particular, if $l > l^{u*}$, increases in employment do not affect the negotiated wage, so, if $C(l)$ were constant for these levels of employment, the cost of the marginal worker would be just $\overline{w}$, as this marginal worker would be unemployed if organization took place.

Figure 3 displays the labor demand curve (R'), the wage-employment curve ($w(l)$), and the marginal cost curve ($MC_L = (w(l)l)'$). The optimal employment level lies at the point where the marginal revenue and the marginal cost curves intersect.

The possibility that the threat of unionization can lead to overemployment was first proposed by Dickens (1986), and to the best of our knowledge has not been developed further. Using a different modeling strategy, Dickens showed that the threat of unionization may increase employment, because firms may want to overhire to reduce the probability with which each employee can obtain a union job, and thus reduce the wage necessary to avoid unionization. However, since, in his model, the choice of employment does not influence the set of bargaining outcomes, he finds that the opposite result is also possible.

Interestingly, Stole and Zwiebel (1996) and (1996b) have also shown that overemployment

will result from a model in which, due the presence of quasi-rents, wages are determined through individual bargaining between the firm and its workers. In that case, as in this section, hiring above the competitive level allows the employer to strike a better wage deal.

The model assumes that no payments are made in the first period. This assumption rules out the possibility that firms require workers to pay a fee to be hired, or, more realistically, to accept low wages while they are unskilled. This would allow firms to recover the rents they are forced to grant to workers once they acquire the ability to unionize. If firms were able to fully recover these rents ex ante, the threat of unionization would not lead to a departure from the competitive outcome. However, it is unlikely that firms can do that since, on the one hand, the difference between the union wage and the competitive wage is substantial (in the range of $15 - 20\%$), and, on the other hand, union workers appear to have relatively long tenures. If workers face borrowing constraints, or there are statutory minimum wages, firms may not be able to reduce initial wages sufficiently to compensate for all the rents they will have to grant workers when the latter acquire firm-specific skills.

III Union Formation under Asymmetric Information

According to Proposition 1, unionization takes place only when it increases the surplus associated with the relationship between the firm and its workers. This result was predicated on the assumption that management and workers have the same information about the expected outcome of unionization. This assumption, however, is problematic. First, it may not be representative of the situation in which most workers find themselves when deciding whether to unionize. One of the main roles of managers is to evaluate the profitability of different courses of action, and their own choices greatly influence future profitability. Therefore, it is likely that managers can form a better assessment of the firm's profit prospects than workers, who lack managers' time and skills and their access to profit-relevant information. Second, assuming symmetric information may be essential to rule out inefficient unionization: as the literature on bargaining with asymmetric information shows, the presence of asymmetric information may lead to inefficient outcomes, which would not be reached if all parties had the same information. Going back to the interpretation of Proposition 1 in terms of the Coase Theorem, assuming complete information in this framework basically amounts to assuming that the firm and its workers will agree to an efficient outcome.

Therefore, this section analyzes the firm's wage and employment determination problem

when employers are better informed than workers about the expected outcome of unionization.[12] To simplify, we assume that firms can be of only two types, high-value (H), and low-value (L) firms, with $R_H > R_L$, and $R'_H(l) > R'_L(l)$, for all $l \geq 0$. Employers know their firm type, but workers only know the distribution of types, which is common knowledge and given by q, the probability that a firm is high-value.

It follows that when employed workers have to decide whether to accept the offered wage, w, or organize, they do not know the wage and employment that would result from unionization. Based on the distribution of types and the wage offer made by the firm, workers update their beliefs about the firm's type and decide whether to accept the wage offer or to organize a union.

If a union formed, the bargaining outcome would depend on the union's assessment of the firm's type. Since union officials typically have time and skills to assess firm performance not available to individual workers, it is reasonable to assume that they will learn additional information with which to update workers' beliefs. We will assume that, in fact, the union is able to observe the firm's type. This assumption is made only for simplicity. All that is needed for the results is that either the union observes some noisy signal of the firm's type, or, even if it has the same information than individual workers, that the outcome of the asymmetric information bargaining game between the firm and the union involves some information revelation (that is, some separation).[13] The result of union-management bargaining is thus $w^u_H(l)$ in high-value firms and $w^u_L(l)$ in low-value firms, where $w^u_H \geq w^u_L$.

To see how the presence of asymmetric information may alter the results in the previous section, let us abstract again for a moment from employment determination, and assume, as in section II, that employment is given by l and that the wage in case of unionization is w^u_i, where i stands for the firm type. If both firms offered the same wage, the wage offer would not reveal any information, and workers' expected payoff from unionization would be $qw^u_H + (1-q)w^u_L - C$. If this expected payoff is not too high, L-firms would find optimal to act as H-firms, and avoid unionization as in the previous section. However, if q or $w^u_H - w^u_L$ is high enough, the expected wage from unionization will be high, making it very costly for L-firms to avoid organization. L-firms would try to signal their low value to workers, but they cannot credibly do so while at the same time avoiding organization, as, in that case, it would clearly be optimal for H-firms

[12] One could be interested in analyzing the twin model in which workers have private information about their outside wage or cost of organization. Although such a model could lead to inefficient unionization if firms miscalculate the wage required to convince workers not to organize, it is hard to imagine how it could explain changes in the probability of organization or the relationship between this probability and other variables.

[13] In fact, most results –all involving separating equilibria– would still hold, even if none of these requirements held.

to imitate them. Therefore, if the ex ante expected wage from organization is high enough, the only alternative for L-firms is to trigger organization, thus revealing their type to the union. H-firms will still find it profitable to avoid the risk of unionization, so that only L-firms become unionized. This argument thus suggests that asymmetric information may lead to union formation, even if unions do not increase rents as required by Proposition 1.

If employment is not fixed, however, L-firms may have another instrument to signal their type: if labor is less productive in those firms, they can reduce employment to dissuade H-firms, for which the reduction would be more costly, from imitating them. If L-firms found it profitable to signal their type in this way, then some firms may hire less workers than predicted by the complete information model. Moreover, the availability of a signal other than unionization may also imply that inefficient unionization is not possible even under asymmetric information.

To analyze the firm's problem under asymmetric information formally, let $q(w,l) = \Pr(H|w,l)$ represent the probability workers assign to the firm being high-value given that a wage w has been offered and l workers have been hired, and $o(w,l)$ the probability with which workers organize given (w,l).

In order to determine its optimal strategy, the firm needs to know workers' expected payoff from unionization. This expected payoff will depend on workers' beliefs about the firm's type induced by the wage offer and on the expected bargaining outcome at each type of firm. Therefore, we first need to characterize these bargaining outcomes. The following lemma shows that, as one would expect, unions can extract larger rents from H-firms.

Lemma 2 *Let $(l_i^u(l), w_i^u(l))$ $(i \in \{L,H\})$ be the bargaining outcome at a firm of type i for a given l. Then*

$$l_H^u(l)(w_H^u(l) - \overline{w}) \geq l_L^u(l)(w_L^u(l) - \overline{w}) \tag{3}$$

Therefore, if we let $w_i(l) = \frac{l_i^u(l)}{l}(w_i^u(l) - \overline{w}) + \overline{w} - \frac{C(l)}{l}$ represent the payoff that workers would expect to obtain if they organized at a firm of type i, Lemma 2 implies that $w_L(l) \leq w_H(l)$. Moreover, let $w(l) = q(w,l)w_H(l) + (1 - q(w,l))w_L(l) - \frac{C(l)}{l}$ be the payoff from unionization expected by workers with beliefs given by $q(w,l)$. Then, $w_L(l) \leq w(l) \leq w_H(l)$ for any $q(w,l)$.

The wage and employment determination problem is a dynamic game with asymmetric information. To reduce potential multiplicity of equilibria, we will use a refinement of the Perfect Bayesian Equilibrium as our equilibrium concept. An equilibrium in our model will be a Perfect Bayesian Equilibrium that satisfies the *Intuitive Criterion* (Cho and Kreps (1987)).

This criterion requires out-of-equilibrium beliefs to assign probability zero to players choosing strategies that are certain to yield a payoff lower than their equilibrium payoff and that, on the other hand, could be preferred by other players to their respective equilibrium payoffs.[14]

The following proposition shows that there are no pooling equilibria.

Proposition 3 *There are no pooling equilibria.*

The intuition behind this result is as follows. At a pooling equilibrium, the pooling wage, w_p, would be greater than $w_L(l_p)$, where l_p is the pooling level of employment. Therefore, it would be possible to increase profits by decreasing employment if, by doing so, workers were led to believe that the firm is of type L. Now, the fact that $R'_H > R'_L$ implies that a given reduction in employment lowers H's revenues more than L's revenues. Thus, it would be possible to find an $l < l_p$ such that H preferred the equilibrium play to $(w_L(l), l)$, while L preferred $(w_L(l), l)$ to (w_p, l_p), which would violate the Intuitive Criterion.

Let (w_i^*, l_i^*) be the choice of a firm of type i under complete information, as computed in Proposition 2. Now, note that H-firms can always do better by choosing (w_H^*, l_H^*) rather than any other (w, l) that revealed their type. Therefore, as the next proposition shows, at any equilibrium in which firms reveal their type, H-firms will set (w_H^*, l_H^*).

Proposition 4 *At any separating equilibrium, H-firms avoid organization with certainty and set* $(w_H, l_H) = (w_H^*, l_H^*)$.

To see what L-firms will do in equilibrium, let $\bar{l} < l_H^*$ be defined by: $R_H(\bar{l}) - w_L(\bar{l})\bar{l} = R_H(l_H^*) - w_H^* l_H^*$. That is, $\bar{l}$ is the level of employment such that a firm of type H would be indifferent between hiring $\bar{l}$ workers and passing as an L-firm, and selecting its complete information optimum. For $l < \bar{l}$, H-firms prefer their complete information optimum, while for $l > \bar{l}$ they prefer to pass as L-firms.

Proposition 5 *If* $\bar{l} > l_L^*$, *then at any separating equilibrium:* $l_L = l_L^*$, *and organization is avoided with certainty.*

The interpretation of the proposition is straightforward. If the differences between firms are such that H-firms could not possibly find it profitable to imitate the optimal behavior of L-firms, then the presence of asymmetric information is irrelevant: each firm behaves exactly as if there were no informational asymmetries. In this case, the overemployment and union

[14] All results below regarding separating equilibria would still hold if we imposed the weaker requirement that workers do not assign a positive probability to firms playing strictly dominated strategies.

avoidance results found for the perfect information case would extend to all types of firms under asymmetric information.

If $\bar{l} \leq l_L^*$, however, H-firms would like to imitate L-firms if the latter were behaving as in the complete information case. Therefore, L-firms will alter their behavior to reveal their type to workers. Because the marginal productivity of labor is lower in these firms, they may signal their type by reducing employment to levels for which imitation would not be profitable for H-firms. On the other hand, since it is also less costly for L-firms to endure unionization, they may trigger organization to reveal their type, as in the previous section. The following propositions show what happens in this case.

Proposition 6 *If* $\bar{l} \leq l_L^*$, *and*

$$R_L(\bar{l}) - w_L(\bar{l})\bar{l} \geq R_L(l_L^{u*}) - w_L^{u*} l_L^{u*}, \tag{4}$$

i. The unique separating equilibria at which $o(w_L, l_L) = 0$ *satisfies* $(w_L, l_L) = \big(w_L(\bar{l}), \bar{l}\big)$.

ii. There exist other separating equilibria with $o(w_L, l_L) > 0$, *and* $l_L > \bar{l}$.

Proposition 7 *If* $\bar{l} \leq l_L^*$, *and*

$$R_L(\bar{l}) - w_L(\bar{l})\bar{l} < R_L(l_L^{u*}) - w_L^{u*} l_L^{u*}, \tag{5}$$

i. There do not exist separating equilibria at which $o(w_L, l_L) = 0$.

ii. There exist separating equilibria at which $o(w_L, l_L) \in (0, 1]$.

If condition (4) holds, L-firms will signal their type by reducing employment rather than by becoming organized.[15] Therefore, the overemployment effect of the threat of unionization will be reduced in L-firms and could even be reversed. If $l_L^{u*} < l_{Lc}$ (where l_{Lc} is the level of employment that would result if L-firms acquired their labor input in a perfectly competitive labor market), then it is possible that $l_L < l_{Lc}$. However, there is a limit to this reduction in employment, as condition (4) ensures that $\bar{l}$ cannot be much lower than l_L^{u*}. (If organization costs were zero, condition (4) would guarantee that $l_L \geq l_L^{u*}$.) On the other hand, if condition (5) holds, the level of employment necessary to avoid being imitated by H-firms is so low that

[15] Even if condition (4) holds, equilibria with a positive probability of organization may exist, in which L-firms use both signals at once. However, equilibria of this form have the property that workers' beliefs are not monotonic: they require workers to believe that a firm is of type L if it offers $(w_L(l_L), l_L)$, while believing that any other (w, l) such that $w = w_L(l)$ and $l \in [\bar{l}, l_L)$ could have been chosen by a firm of type H. This feature makes them less compelling predictions than the pure strategy equilibria.

L-firms prefer to trigger unionization to signal their type. Therefore, propositions 4-7 imply that *less profitable firms are more likely to become unionized*, as suggested by the discussion at the beginning of the section.

IV Implications and Empirical Evidence

a Determinants of Unionization

It is often argued that workers will be more likely to join unions in those firms or industries earning larger rents, as the demand for union services should be greater where the potential gains from organization are larger. This would be true if firms did not act at all to prevent organization, or if their actions were unrelated to rent size. However, large gains to workers mean large losses to the firm. Therefore, it is reasonable to expect profitable employers to resist union organization more strongly. Proposition 1 shows that, in the absence of asymmetries of information, all firms, independently of their profitability, find it optimal to avoid unionization, so that there are no a priori reasons to believe that more profitable firms are more likely to become unionized. Moreover, propositions 4-7 show that, under asymmetric information, the probability of unionization within a group of ex ante identical firms is negatively related to the size of the (ex post) rents to be divided between the firm and its employees. Put differently, unionization is bad news even if it has a negligible direct effect on profits. Ruback and Zimmerman (1984) provide an ideal test of this prediction in a study written to estimate the effect of unions on profits. Using a sample of large publicly traded firms, the authors computed abnormal stock returns for the month in which the union petitions for an election and the month in which the election outcome is certified. If the election petition conveyed no information about the firm's type, the stock market price of a firm where a union election is petitioned would decrease at the date of the petition (since a petition is associated with a greater probability of organization), but would go up, at least partially, if the union loses the election. If, however, there is a signaling effect, the positive impact of a union loss on the stock price should be smaller, since, although the union loss implies higher future profits (driving the price up), the negative information conveyed by the unionization attempt implies that profit expectations are lower after the union loss than before the election petition. Ruback and Zimmerman report that, in the month of an election petition, abnormal returns are negative and statistically significant both in those firms in which the unionization attempt will end up being successful and in those where it will lead to a union loss. Although the negative abnormal

returns are smaller in the latter group of firms, the difference is not statistically significant. In the month in which the election outcome is certified, abnormal returns are negative in firms where the union wins *and* in those in which the union loses, although in both cases they are not significantly different from zero. That is, the news of a union loss does not significantly increase profit expectations. This result suggests that investors may interpret the inability of a firm to avert unionization attempts as bad news about the firm's prospects irrespective of whether unionization ends up occurring or not, as implied by the model in this paper.[16]

According to the model, no prediction can be made about the relationship between ex ante expected rents and unionization, since the probability of organization does not depend on the location, but on the shape of the distribution of rents. Only if rents are so low (and consequently union wages too) that unionization is ex ante unprofitable for workers independently of the wage offer, will knowledge of ex ante expected rents be informative about the probability of organization. A fall in rents known to to make unionization unprofitable for any wage is certain to bring the probability of organization down to zero. Using different data sets and methodologies, Bronars and Deere (1993) and Bronfenbrenner (1997) find either no or a negative correlation between firm growth or profitability and unionization. Although these results do not provide conclusive evidence for or against the predictions of the model, they certainly cast doubt on the common view that unions will organize more profitable establishments.

The model also predicts that unionization will be more likely the more probable it is that unions increase surplus (either by improving productive efficiency or providing direct utility to their members). Duncan and Stafford (1980) provide evidence that unionization is associated with working conditions "which require interdependent worker behavior as part of the production process"(*ibid.* p. 369), such as working with machines or inflexible hours of work. In the authors' interpretation, these working conditions make collective representation efficient, thus supporting the relevance of efficiency explanations of union incidence. However, the evidence on the productivity effects of unions is mixed (Booth (1995)).

Two other factors have been shown to affect the probability of unionization. First, increases in costs of organization will have the expected effect of making unionization more difficult (as smaller costs make it more likely for inequalities (2) in Proposition 1, and (5) in Proposition 7 to hold). Second, the degree of uncertainty about the firm's profits will influence the probability

[16]It could be argued that this may be due to investors' anticipation of the election result. But the fact that those firms where the union will lose the election experience significant negative abnormal returns the month of the election petition (and not significantly different from the abnormal returns of firms in which the union will win) suggests that investors' anticipation is not complete.

of unionization, although this effect cannot be signed a priori. If firms are very different, so that l_L^* is much smaller than l_H^*, then asymmetric information will probably not imply any changes with respect to the complete information outcomes. If firms are very similar, it is not likely that condition (5) in Proposition 7 holds, and therefore unionization would also be rare. Therefore, unionization will tend to be more likely when the difference between firms is substantial but not extreme.

b Union Effects on Profits

If unionization is due to efficiency enhancing effects of unions, it is not possible to say a priori whether unionized firms will tend to be more or less profitable than nonunion ones. The asymmetric information model, however, implies that profits in union firms will be lower than in nonunion firms. Moreover, it also shows that unionization is likely to be related to unobservable firm characteristics that affect the firm's ability to generate profits. Therefore, if we want to measure the direct effect of unions on profits, simply comparing the profits of union and nonunion firms will be misleading. Conventional OLS estimates of the effect of unions on profits will thus be biased. Moreover, since unionization is more likely in firms with (unobservably) worse profit prospects, those estimates will overestimate the effect of unions on profits. The use of longitudinal data may not ameliorate the problem, as changes in union status are likely to be caused by changes in unobserved profitability. The role of organization costs in the determination of unionization, however, suggests that finding an instrument that is likely to be correlated with the costs of organization (e.g. state laws or demographic or attitudinal characteristics) but not with firm profitability could be a way to control for the above bias.

c Union Effects on Nonunion Wages and the Union-Nonunion Wage Gap

Although it is not possible to obtain precise comparative statics results without imposing specific functional forms, the model does not provide support for the argument that a greater wage gap should be associated with an increased probability of organization. On the contrary, since increases in the cost of organization are likely to reduce the probability of unionization and to increase the wage gap (by reducing the wage paid by nonunion firms), one should expect that, if anything, unionization and the wage gap will tend to move in opposite directions. It should be noted, however, that all results in this paper refer to the behavior of a single firm, which takes the outside wage $\overline{w}$ as given. If an industry were composed of (ex ante) identical firms, and labor supply were infinitely elastic at wage $\overline{w}$, the results of the model would extend

automatically to that industry's partial equilibrium, but potential general equilibrium effects have not been considered.

With this caveat in mind, the model's predictions appear to be consistent with the U.S. experience. As Figure 1 shows, union density has exhibited a large decline since the 1970s, while the union wage premium seems to have remained constant, or even trended upwards. In fact, Freeman (1986) finds a negative correlation between the aggregate wage gap and new unionization, which also appears in more disaggregated studies (Linneman and Wachter (1986), Linneman, Wachter, and Carter (1990)). Data from 1930-1974, although less reliable than more recent micro-data, also show that those periods in which the greatest increases in union density took place were accompanied by lower union wage gaps than periods of union decline or moderate growth.[17] The cross-sectional pattern of union density and the union wage premium in the U.S. also hints at a negative correlation between the two variables when comparing regions (the U.S. South vs. other regions, or rural vs. urban areas), sectors (the private vs. the public sector), or establishments of different size (with smaller establishments being characterized by a smaller union density and a higher wage gap), at least for manufacturing.[18] Interestingly, the same pattern seems to hold when comparing different countries, as the comparison of the U.S. with relatively similar Canada exemplifies: the higher union density in Canada is not accompanied by higher wage premia.

As in the case of profits, it is worth stressing that the model predicts that OLS estimates of the effect of unionization on union wages will be biased due to firm selection into the union sector. In this case, the bias will have a negative sign, as high-wage firms (H-firms) are more likely to remain nonunion.[19]

The model also predicts that the threat of unionization will lead nonunion firms to increase their wages. This prediction is supported by interview evidence (Rees and Schultz (1970), Foulkes (1980)), and, less conclusively, by econometric studies, which tend to find a positive effect of union density on nonunion wages. Insofar as the variation in union density is caused

[17] Pencavel (1991), pp. 26-30.

[18] This description is based on Lewis's comprehensive survey of empirical measures of the wage gap (Lewis (1986)). We are not aware of any more recent study showing correlations of a different sign. Note that we have not presented any result on how the wage gap varies with the extent of unionism in an industry, which would be a more direct test of the theory. The reason is that results seem to be extremely sensitive to the choice of the variable measuring the extent of unionism, and seem to pick up the effects of omitted variables more than the effects of unionism. For a discussion, see Lewis (1986).

[19] If we cannot control for all observable firm characteristics, this bias could be reduced or even reversed if, in an industry in which unionization takes place with positive probability, there are firms with rents so low that unionization is not ex ante profitable for any wage offer. In this case, these low-wage firms will be overrepresented within the nonunion firms. Of course, there are other sources of bias associated with unobservable worker or job characteristics that have not been contemplated in this paper (Card (1996)).

by differences in organization costs, the model in this paper would yield the same prediction. However, estimates of the effect of union density on nonunion wages are generally based on cross-sectional regressions, and thus need to be interpreted with care: if unionization were more likely in industries that would have paid higher wages in any case, cross-sectional estimates could just be capturing this fact (Neumark and Wachter (1995)).

A drawback of the research that uses union density as a regressor is that, even if it avoids the estimation biases just mentioned, we can only be sure that it is measuring the effect of union density on nonunion wages. Whether it is also measuring the effect of the threat of unionization (the type of effect to which both this paper and the standard union threat view refer) is, however, arguable, since it is not clear that union density is a good proxy for the strength of the union threat. In the study that most closely approximates the ideal experiment to test the model's predictions, Ichniowski, Freeman, and Lauer (1989) trace the effects of changes in state collective bargaining laws on union and nonunion police compensation. They find that changes in bargaining laws that make it easier for public sector employees to bargain collectively increase union density and increase the compensation of both union and nonunion police. Unfortunately, these results refer only to a particular class of public sector workers, so that no general conclusions can be extracted.

d Cyclical Implications

The model suggests potential cyclical implications of the threat of unionization in an uncertain environment. If uncertainty has an aggregate component, and managers have better information about the realization of uncertainty, an adverse shock to firms' profitability may lead to an increase in union organizing activity, as many firms are forced to reveal their bad profit prospects in this way.[20] Other firms, however, will lay off workers to convince their workforce that a negative aggregate shock has indeed occurred.[21] The threat of unionization may thus lead to greater employment responses to shocks than those implied by a competitive labor market, a possibility that should be explored further.

[20] The model could thus provide an explanation for the finding by Bronars and Deere (1993) that GNP growth was significantly above average before periods of union recognition activity and significantly below average after those periods.

[21] This possibility has been analyzed in the context of implicit insurance contracts (Grossman, Hart, and Maskin (1983)), and in bargaining models (Acemoglu (1995) and Kennan (2002)).

V Suggestions for Further Research

The analysis in this paper provides a first step in the understanding of the operation of partially unionized labor markets and the determinants of unionization. This section outlines some directions along which this research can be fruitfully extended.

First, integrating the model within a general equilibrium framework, with capital investment and sectors with different exposure to the threat of unionization, would allow us to determine the net effect of workers' ability to unionize on equilibrium wages, aggregate employment, and union membership. The latter may be especially important because several variables in the model are likely to depend on aggregate union membership. For example, costs of organization may depend on the number of unionized workers in an industry. Moreover, we must take these externalities into consideration if we are to compare the performance of different systems of collective bargaining (for example, decentralized versus industry-level bargaining), as one of the main dimensions along which these systems vary is the degree to which labor market participants internalize organization and bargaining externalities. Although the model indicates that increases in union density are likely to be associated with an improvement of the welfare of nonunion workers, the analysis of market equilibrium is also necessary to assess possible equilibrium effects that may alter this prediction. Finally, since it is difficult to obtain reliable measures below the industry level for most of the variables of interest, aggregate predictions should also be easier to bring to the data.

Second, an explicitly dynamic extension of the model would clarify the model's cyclical implications, and allow us to obtain predictions that can be more readily tested using time-series data. It would also shed light on the factors behind the development of unionism in the U.S. private sector, characterized by a rapid rise in union density in the period between the Great Depression and the mid 1950s, and by a monotonic decline ever since. To understand this development, the model in this paper indicates that special attention should be paid to the evolution of the costs of organization to workers.

Third, the model offers a host of predictions and suggests empirical strategies. Work specifically aimed at testing these predictions and those generated by the extensions outlined above, is required. In particular, it would be especially fruitful to obtain reliable measures of the costs of organization to workers and their relation to wages, employment and union membership.

Lastly, more sophisticated contract forms should also be considered. The model restricts contracts to be noncontingent, while suggesting that efficiency gains could be possible if com-

pensation were made to depend on potentially verifiable variables like profits or employment. Similarly, implicit contracts which made compensation contingent on firm profitability could be, in principle, sustainable if firms and workers interact repeatedly. Studying possible private responses to the inefficiency caused by asymmetric information is also necessary to evaluate the potential for efficiency-enhancing policy intervention. In particular, it would be worth investigating whether measures aimed at reducing the informational asymmetries between firms and workers, such as, say, compulsory worker representation on company boards (even if lacking any decision rights) or information disclosure requirements could alleviate this inefficiency.

VI Conclusion

Nonunion employers in a partially unionized industry are not likely to behave as described by standard competitive models. Rather than taking wages as given, they will set wages and employment strategically to ensure that their workers do not organize. In this paper, we have shown that, as a result, labor market outcomes will depart significantly from those derived from models in which the nonunion sector is assumed to operate competitively.

In particular, nonunion wages will be set above the competitive level to eliminate workers' incentives to organize. Despite this increase in the average cost of labor, employment will generally not fall, as nonunion firms have the incentive to overemploy to reduce the average rents that workers can expect to obtain if they organize. If employers are better informed than workers about profitability, however, less profitable firms may have an even stronger incentive to lay off workers if, by doing so, they can convince the remaining employees of their bad profit outlook, thus persuading them to moderate their wage demands. Although its macroeconomic consequences are yet to be derived explicitly, the model suggests that the effect on employment of negative shocks could be amplified by firms' recourse to layoffs as a device to lower wage demands and avoid unionization.

By explicitly modeling employer behavior towards unionization, this paper has also provided a theory of the incidence of unionism. The interest of such a theory is twofold. On the one hand, it offers a new framework to study the causes and consequences of the decline of U.S. unionism in the last two decades. Moreover, unlike one-sided theories, which take either wages or union density as given, this framework is able to offer predictions about the joint behavior of these variables. Of special interest is the prediction that union density and the union wage premium will tend to be negatively correlated, which seems to fit the pattern observed in the

U.S.. This pattern is, however, difficult to rationalize by theories of union formation that take the wage premium as given.

On the other hand, a theory of the simultaneous determination of unionization, wages and employment addresses the key issue in the empirical literature examining the effects of labor unions, namely the nonrandom determination of union status. According to the model, the probability of union organization does not depend on observable firm characteristics not directly related to the costs of organization to workers, or the potential effect of unions on total rents. It follows that observable measures of profitability are not likely to be informative to predict union organization. In contrast, unobservable profit prospects play a large role in the determination of union status. That is, unionization is likely to be correlated with unobservable firm characteristics, which, at the same time, have a direct effect on firm performance. Therefore, the model predicts that OLS estimates of the effects of unionization will be biased. Against the often held assumption that unions will organize the (unobservably) more profitable firms, the model predicts that organization will be more likely in (unobservably) less profitable firms, and thus pins down the sign of the estimation bias. Apart from cautioning us about the existence and sign of this bias, the model also has implications concerning possible remedies. In particular, it does not favor the use of longitudinal data to identify union effects, as fixed firm characteristics, as long as they are observable by workers, will be unrelated to changes in union status. The role of organization costs in the determination of union status, however, suggests that factors unrelated to firm performance, but correlated with these costs (like changes in the legal environment regulating collective bargaining, or differences in workers' attitudes towards unions), may be used as instruments for unionization.

VII Appendix

Proof of Lemma 1. Let problem (A) be: $\max_{l^u} U_A(l^u) = l^u(R^{u'}(l^u) - \overline{w})$. If U_A is strictly concave, then problem (A) has at most one solution, l^{u*}, and, assuming $R^{u'}(0) > \overline{w}$, and $R^{u'}(l) < \overline{w}$ for l large enough, $l^{u*} > 0$ is given by $R^{u'}(l^{u*}) = \overline{w} + l^{u*}R^{u''}(l^{u*}) > \overline{w}$. Therefore, if l^u_c is such that $R^{u'}(l^u_c) = \overline{w}$, then $l^{u*} < l^u_c$. Since $U''_A = 2R^{u''} + l^u R^{u'''}$, concavity of $R^{u'}$ guarantees concavity of U_A. Let $w^{u*} = R^{u'}(l^{u*}) > \overline{w}$.

Now, the union solves:

$$\begin{aligned} \max_{l^u, w^u} \quad & l^u(w^u - \overline{w}) \\ s.t. \quad & l^u = \arg\max\{R^u(l^u) - w^u l^u, \ for \ l^u \leq l\} \end{aligned}$$

It cannot be the case that $w^u < R^{u'}(l)$, since that would lead to $l^u = l$, and the union would be better off setting $w^u = R^{u'}(l)$, since that would increase the wage without lowering employment. Therefore $w^u \geq R^{u'}(l)$ for all l, which implies that l^u will satisfy $R^{u'}(l^u) = w^u$. The problem of the union thus becomes:

$$\begin{aligned} \max_{l^u} \quad & l^u(R^{u'}(l^u) - \overline{w}) \\ s.t. \quad & l^u \leq l \end{aligned}$$

If l^{u*} is feasible, then it maximizes this problem. Thus $l^u(l) = l^{u*}$ for $l \geq l^{u*}$. Now, concavity of U_A ensures that $R^{u'}(l^u) - \overline{w} - l^u R^{u''}(l^u) > 0$ for $l^u < l^{u*}$. Therefore, for $l < l^{u*}$, $l^u(l) = l$. This proves 2) and 1) (since we have already shown that $l^{u*} < l^u_c$). 3) follows from 2), 1) and the fact that $w^u = R^{u'}(l)$. Finally, 4) follows from 1) - 3), as, for $l < l^{u*}$, $R^u(l^u(l)) - l^u(l)w^u(l) = R^u(l) - lR^{u'}(l)$, which is increasing in l due to the concavity of R; and, for $l \geq l^{u*}$, $R^u(l^u(l)) - l^u(l)w^u(l) = R^u(l^{u*}) - l^{u*}R^{u'}(l^{u*})$. ■

Proof of Proposition 1. If the firm offers a wage below $w(l)$, workers form a union and profits are $\Pi_U(l) = R^u(l^u(l)) - w^u(l)l^u(l)$. If it offers $w(l)$, workers accept and profits are $\Pi(l) = R(l) - l^u(l)w^u(l) + (l - l^u(l))\overline{w} + C(l)$. From Lemma 1, we know that $\Pi_U(l^{u*}) = R^u(l^{u*}) - w^{u*}l^{u*} \geq \Pi_U(l)$, for any l. Now, at the optimal choice of l, $\Pi(l^*) \geq \Pi(l^{u*}) = R(l^{u*}) - w^{u*}l^{u*} + C(l^{u*})$. It follows that a sufficient condition for optimality of union avoidance is that $R^u(l^{u*}) \leq R(l^{u*}) + C(l^{u*})$. The opposite inequality is necessary but not sufficient for unionization, since, if $l^* > l^{u*}$, it can be the case that $R^u(l^{u*}) - C(l^{u*}) > R(l^{u*})$, yet $R^u(l^{u*}) - C(l^*) < R(l^*) + (l^* - l^{u*})\overline{w}$,

which would imply $\Pi(l^*) = R(l^*) - w^{u*}l^{u*} + (l^* - l^{u*})\overline{w} + C(l^*) > \Pi_U(l^{u*})$. ■

Proof of Proposition 2. Proposition 1 implies that it is optimal for the firm to avoid organization. If the firm sets employment l and avoids unionization, profits are given by $\Pi(l) = R(l) - l^u(l)w^u(l) + (l - l^u(l))\overline{w} + C(l)$. For $l < l^{u*}$, $l^u(l) = l$, so that profits are $R(l^u(l)) - w^u(l)l^u(l) + C(l)$. But from Lemma 1 we know that $R(l^u(l)) - w^u(l)l^u(l)$ is increasing in l for $l < l^{u*}$. If $C(l)$ is nondecreasing that implies that $l^* \geq l^{u*}$. Now, for $l > l^{u*}$, $l^u(l) = l^{u*}$, and $w^u(l) = w^{u*}$, so that $\Pi'(l) = R'(l) - \overline{w} + C'(l)$. Hence, for $l < l_c$, $\Pi'(l) > 0$, and $l^* \geq l_c$. If $C'(l_c) = 0$, then $\Pi'(l_c) = 0$ and $l^* = l_c$. If $C'(l_c) > 0$, then $\Pi'(l_c) > 0$, and $l^* > l_c$. ■

Proof of lemma 2. The fact that $R'_H > R'_L$ implies that there exists a pair (w^u, l^u) on the R'_H curve such that $l \geq l^u \geq l^u_L(l)$, and $w^u > w^u_L(l)$, so that $l^u(w^u - \overline{w}) > l^u_L(l)(w^u_L(l) - \overline{w})$. Therefore, since $(l^u_H(l), w^u_H(l))$ is preferred to any other (w^u, l^u) on R'_H, $l^u_H(l)(w^u_H(l) - \overline{w}) \geq l^u(w^u - \overline{w}) > l^u_L(l)(w^u_L(l) - \overline{w})$. ■

Proof of proposition 3. Let $(w_L, l_L) = (w_H, l_H) = (w_p, l_p)$, and suppose $o(w_p, l_p) = 0$. Now, define $l_e < l_p$ by $R_H(l_p) - w_p l_p = R_H(l_e) - w_L(l_e)l_e$, so that, if $l < l_e$, H[22] prefers the equilibrium play to $(w(l_e), l_e)$ for any $q(w(l_e), l_e)$. Therefore, for the equilibrium to satisfy the Intuitive Criterion, it has to be the case that if $l < l_e$, then $q(w_L(l), l) = 0$. But, in this case:

$$\begin{aligned} R_L(l_e) - w_L(l_e)l_e - (R_L(l_p) - w_p l_p) = \\ R_L(l_e) - (R_H(l_e) + w_p l_p - R_H(l_p)) - (R_L(l_p) - w_p l_p) = \\ (R_H(l_p) - R_L(l_p)) - (R_H(l_e) - R_L(l_e)) > 0, \end{aligned}$$

where the last inequality follows from $R'_H > R'_L$ and $l_e < l_p$. Therefore, there exists a profitable deviation for L, and $(w_L, l_L) = (w_H, l_H) = (w_p, l_p)$ cannot be an equilibrium with $o(w_p, l_p) = 0$. The same argument would apply for $o(w_p, l_p) > 0$, and for potential partly revealing equilibria.■

Proof of proposition 4. Let $(l_H, w_H) \neq (l^*_H, w^*_H)$. Since l^*_H is the unique optimal l under perfect information, $R_H(l^*_H) - w_H(l^*_H)l^*_H > R_H(l^u_H(l)) - w^u_H(l)l^u_H(l)$, and $R_H(l^*_H) - w_H(l^*_H)l^*_H > R_H(l) - w_H(l)l$ for all l. Moreover, for any q, $R_H(l^*_H) - w(l^*_H)l^*_H \geq R_H(l^*_H) - w_H(l^*_H)l^*_H$. Therefore, $l^*_H, w^*_H + \epsilon$, for $\epsilon \geq 0$ small enough, is a profitable deviation, and (l_H, w_H) cannot

[22] For brevity, H and L denote H- and L-firms, respectively.

be H's equilibrium play. The same argument would apply if H randomized at the proposed equilibrium play. ■

Proof of proposition 5. Let $l_M \equiv arg\max_l R_L(l) - w_H(l)l$. Then, the following are necessary conditions for the existence of a separating equilibrium:

$$\begin{aligned} R_H(l_H) - w_H(l_H)l_H &\geq o(w_L, l_L)(R_H(l_H^u(l_L)) - w_H^u(l_L)l_H^u(l_L)) \\ &+ (1 - o(w_L, l_L))(R_H(l_L) - w_L l_L) \end{aligned} \tag{6}$$

$$\begin{aligned} o(w_L, l_L)(R_L(l_L^u(l_L)) - w_L^u(l_L)l_L^u(l_L)) + (1 - o(w_L, l_L))(R_L(l_L) - w_L l_L) &\geq \\ R_L(l_M) - w_H(l_M)l_M \end{aligned} \tag{7}$$

$$\begin{aligned} o(w_L, l_L)(R_L(l_L^u(l_L)) - w_L^u(l_L)l_L^u(l_L)) + (1 - o(w_L, l_L))(R_L(l_L) - w_L l_L) &\geq \\ R_L(l_L^{u*}) - w_L^{u*}l_L^{u*} \end{aligned} \tag{8}$$

A condition like (8) for H is not required as it was already shown in Proposition 4 that H prefers no organization.

Now, for $l < l_L^{u*}$, $R'_H(l) - (w_L(l)l)' = R'_H(l) - (R'_L(l) + R''_L(l)l - C'(l)) > 0$, since $R'_H(l) > R'_L(l)$, $R''_L(l) \leq 0$, and $C'(l) \geq 0$. For $l \in (l_L^{u*}, l_H^*)$, $R'_H(l) - (w_L(l)l)' = R'_H(l) - (\overline{w} - C'(l)) > 0$, since $R'_H(l_H^*) - (\overline{w} - C'(l_H^*)) = 0$, $l < l_H^*$, R_H is concave, and C is convex. Therefore, $R_H(l) - w_L(l)l$ is increasing for $l < l_H^*$, which implies that $R_H(l_H^*) - w_H^* l_H^* > R_H(l) - w_L(l)l$ for any $l < \bar{l}$, where $\bar{l}$ was defined in the text. It follows that, if $\bar{l} > l_L^*$, then, at any separating equilibrium, $(w_L, l_L) = (w_L^*, l_L^*)$, and $o(w_L, l_L) = 0$. To check existence, note that $\bar{l} \geq l_L^*$ guarantees that (6) holds, and optimality of l_L^* guarantees that both (7) and (8) hold as well. Lastly, the beliefs $q(w, l) = 0$ for $l \leq \bar{l}$, and $q(w, l) = 1$ for $l > \bar{l}$ support the equilibrium, and satisfy the Intuitive Criterion. No other (w_L, l_L) can be played at an equilibrium that survives the Intuitive Criterion, as this criterion ensures that $q(w_L^*, l_L^*) = 0$, so that (w_L^*, l_L^*) would be a profitable deviation for L. ■

Proof of proposition 6. i. Let $\bar{l} \leq l_L^*$. First note that it has to be the case that

$$R_L(l_M) - w_H(l_M)l_M \leq R_L(\bar{l}) - w_L(\bar{l})\bar{l}. \tag{9}$$

(That is, it is not possible that H would like to imitate L, and L would like to imitate H at the same time). To see this, assume instead that

$$R_L(l_M) - w_H(l_M)l_M > R_L(\bar{l}) - w_L(\bar{l})\bar{l}. \tag{10}$$

This implies that $l_M > \bar{l}$. Otherwise, by 4. in Lemma 1, and the fact that $w_L(l) \leq w_H(l)$, $R(\bar{l}) - w_L(\bar{l})\bar{l} > R(l_M) - w_L(l_M)l_M > R(l_M) - w_H(l_M)l_M$. Now, by definition, $R_H(\bar{l}) - w_L(\bar{l})\bar{l} > R_H(l_M) - w_H(l_M)l_M$, so that $w_L(\bar{l})\bar{l} < R_H(\bar{l}) - (R_H(l_M) - w_H(l_M)l_M)$, which implies $R_H(\bar{l}) - (R_H(l_M) - w_H(l_M)l_M) > R_L(\bar{l}) - (R_L(l_M) - w_H(l_M)l_M)$. Or, rearranging $R_H(\bar{l}) - R_H(l_M) > R_L(\bar{l}) - R_L(l_M)$, but this is possible only if $l_M < \bar{l}$, which contradicts $l_M > \bar{l}$. Therefore (10) cannot hold.

Let condition (4) hold, that is, $R_L(\bar{l}) - w_L(\bar{l})\bar{l} \geq R_L(l_L^{u*}) - w_L^{u*}l_L^{u*}$. Then, there exist separating equilibria that survive the Intuitive Criterion such that $(w_L, l_L) = \left(w_L(\bar{l}), \bar{l}\right)$, and $o\left(w_L(\bar{l}), \bar{l}\right) = 0$. To see this note that (6) holds by definition of $\bar{l}$, and (9) guarantees that (7) also holds. The same beliefs as in the previous proof would support the equilibrium. Note that in this case, there exist no other equilibria with zero probability of organization, as any equilibrium with $l_L < \bar{l}$ would not survive the Intuitive Criterion, and if $(w_L, l_L) = (w_L(l), l)$ for $l > \bar{l}$, (6) would not hold.

ii. Equilibria with positive probability of organization can exist for $l \geq \bar{l}$ (if $l < \bar{l}$ they would not survive the Intuitive Criterion). To see this, notice that if $R_L(l_M) - w_H(l_M)l_M < R_L(\bar{l}) - w_L(\bar{l})\bar{l}$, and $R_L(l_L^{u*}) - w_L^{u*}l_L^{u*} < R_L(\bar{l}) - w_L(\bar{l})\bar{l}$, then it is possible to find an $l_L = \bar{l} + \epsilon$, with $\epsilon > 0$, and an $o > 0$ such that H prefers (w_H^*, l_H^*) to $(w_L(l_L), l_L)$ if $o(w_L(l_L), l_L) = o$, and L still prefers $(w_L(l_L), l_L)$ to imitating H or becoming organized with certainty. To guarantee that (6) holds, it is enough to set

$$o = \frac{(R_H(\bar{l}+\epsilon) - w_L(\bar{l}+\epsilon)(\bar{l}+\epsilon)) - (R_H(\bar{l}) - w_L(\bar{l})\bar{l})}{(R_H(\bar{l}+\epsilon) - w_L(\bar{l}+\epsilon)(\bar{l}+\epsilon)) - (R_H(l_H^u(\bar{l}+\epsilon)) - w_H^u(\bar{l}+\epsilon)l_H^u(\bar{l}+\epsilon))}$$

For ϵ small enough, (7) and (8) will still hold. ■

Proof of proposition 7. Let $(w_L, l_L) = (w_L(\bar{l}) - t, \bar{l})$, for some $t > 0$, and $o(w_L, l_L) = 1$. Let beliefs be given by $q(w, l) = 1$ for $l > \bar{l}$, and $q(w, l) = 0$ for $l \leq \bar{l}$. These beliefs satisfy the Intuitive Criterion, and, given the beliefs and condition (5), L behaves optimally (as it was shown above that if $\bar{l} < l_L^*$, $R_L(\bar{l}) - w_L(\bar{l})\bar{l} \geq R_L(l_M) - w_H(l_M)l_M$).

Note also that no equilibrium exists at which organization is avoided with certainty, as any

such equilibrium would require $l_L > \bar{l}$, in which case, H would want to imitate L , so that H's incentive compatibility constraint would not hold. There can also exist separating equilibria with $o(w_L, l_L) \in (0, 1)$. Let $\underline{l} < l_L^*$ be defined by $R_L(\underline{l}) - w_L(\underline{l})\underline{l} = R_L(l_L^{u*}) - w_L^{u*} l_L^{u*}$, so that $\underline{l} > \bar{l}$, and let $l_L = \underline{l}$ and $w_L = w_L(\underline{l})$. Then, if $o(w_L, l_L) \geq \frac{R_H(\underline{l}) - w_L(\underline{l})\underline{l} - (R_H(l_H) - w_H l_H)}{R_H(\underline{l}) - w_L(\underline{l})\underline{l} - (R_H(l_H^u(\underline{l})) - w_H^u(\underline{l}) l_H^u(\underline{l}))}$, (6) will hold. On the other hand, since $R_L(\underline{l}) - w_L(\underline{l})\underline{l} = R_L(l_L^{u*}) - w_L^{u*} l_L^{u*}$, L cannot do better by triggering certain organization. Now, the facts that $l_L^* > \underline{l} > \bar{l}$ and $R_L(\bar{l}) - w_L(\bar{l})\bar{l} \geq R_L(l_M) - w_H(l_M) l_M$ guarantee that (7) also holds. Mixed strategy equilibria are also possible for $l > \underline{l}$. However, these equilibria involve nonmonotonic beliefs, as they require $q(w_L, l_L) = 0$ and $q(w_L(l), l) > 0$ for $l \in [\underline{l}, l_L)$. ■

References

Acemoglu, D. (1995). Asymmetric information, bargaining, and employment fluctuations. *International Economic Review 36*(4), 1003–1024.

Blanchflower, D. G. (1997). Changes over time in union relative wage effects in Great Britain and the United States. *NBER Working Paper: 6100*.

Booth, A. (1985). The free rider problem and a social custom model of trade union membership. *Quarterly Journal of Economics 100*, 253–261.

Booth, A. L. (1995). *The Economics of the Trade Union*. Cambridge, U.K.: Cambridge University Press.

Bronars, S. G. and D. R. Deere (1993). Union organizing activity, firm growth and the business cycle. *American Economic Review 83*(1), 203–220.

Bronfenbrenner, K. (1997). The role of union strategies in NLRB certification elections. *Industrial and Labor Relations Review 50*(2), 195–212.

Card, D. (1996). The effect of unions on the structure of wages: A longitudinal analysis. *Econometrica 64*(4), 957–79.

Cho, I.-K. and D. Kreps (1987). Signaling games and stable equilibria. *Quarterly Journal of Economics 102*(2), 179–221.

Corneo, G. (1995). Social custom, management opposition, and trade union membership. *European Economic Review 39*, 275–292.

Dickens, W. T. (1986). Wages, employment, and the threat of collective bargaining by workers. *NBER Working Paper No. 1856*.

Duncan, G. J. and F. P. Stafford (1980). Do union members receive compensating wage differentials? *American Economic Review 70*(3), 355–371.

Ehrenberg, R. G. and R. S. Smith (1997). *Modern Labor Economics. Sixth Ed.* Reading, MA: Addison Wesley.

Farber, H. S. (1990). The decline in unionism in the united states: What can be learned from recent experience. *Journal of Labor Economics 8*(1), S75–S105.

Foulkes, F. K. (1980). *Personnel Policies In Large Nonunion Companies*. Englewood Cliffs, NJ: Prentice Hall Inc.

Freeman, R. B. (1986). The effect of the union wage differential on management opposition and union organizing success. *American Economic Review 76*(2), 92–96.

Freeman, R. B. and M. M. Kleiner (1990). Employer behavior in the face of union organizing drives. *Industrial and Labor Relations Review 43*(4), 351–365.

Grossman, S. J., O. D. Hart, and E. Maskin (1983). Unemployment with observable aggregate shocks. *Journal of Political Economy 91*(6), 907–928.

Hirsch, B. T. (1991). *Labor Unions and the Economic Performance of Firms*. Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.

Hirsch, B. T. and E. J. Schumacher (2000). Earnings imputation and bias in wage gap estimates. *Mimeo*.

Ichniowski, C., R. B. Freeman, and H. Lauer (1989). Collective bargaining laws, threat effects, and the determination of police compensation. *Journal of Labor Economics 7*(2), 191–209.

Kennan, J. (2002). Informational conflict and employment fluctuations. *Mimeo. University of Wisconsin-Madison*.

Lazear, E. (1983). A competitive theory of monopoly unionism. *American Economic Review 73*(4), 631–43.

Lewis, G. H. (1986). *Union Relative Wage Effects: a Survey*. Chicago, IL: University of Chicago Press.

Linneman, P. and M. L. Wachter (1986). Rising union premiums and the declining boundaries among noncompeting groups. *American Economic Review 76*(2), 103–108.

Linneman, P., M. L. Wachter, and W. H. Carter (1990). Evaluating the evidence on union employment and wages. *Industrial and Labor Relations Review 44*(1), 34–53.

Naylor, R. and M. Cripps (1993). An economic theory of the open shop trade union. *European Economic Review 37*, 1599–1620.

Neumark, D. and M. L. Wachter (1995). Union effects on nonunion wages: Evidence from panel data on industries and cities. *Industrial and Labor Relations Review 49*(1), 20–38.

Pencavel, J. H. (1991). *Labor Markets Under Trade Unionism: Employment, Wages and Hours*. Oxford and Cambridge, MA: Basil Blackwell Inc.

Rees, A. and G. P. Schultz (1970). *Workers and Wages in an Urban Labor Market*. Chicago, IL: University of Chicago Press.

Rosen, S. (1969). Trade union power, threat effects and the extent of organisation. *Review of Economic Studies 36*(3), 185–96.

Ruback, R. S. and M. B. Zimmerman (1984). Unionization and profitability: Evidence from the capital market. *Journal of Political Economy 92*(6), 1134–1157.

Segal, I. (2001). Collusion, exclusion and inclusion in random order bargaining. *Review of Economic Studies (forthcoming)*.

Stole, L. and J. Zwiebel (1996). Intra-firm bargaining under non-binding contracts. *Review of Economic Studies 63*(3), 375–410.

Stole, L. and J. Zwiebel (1996b). Organizational design and technology choice under intrafirm bargaining. *American Economic Review 86*(1), 195–222.