

Forthcoming in John T. Addison and Claus Schnabel (eds.) *International Handbook of Trade Unions*, Cheltenham, UK: Edward Elgar, Chapter 4, 2003.

Unions, Bargaining and Strikes

Peter Cramton and Joseph Tracy

August 25, 2002

University of Maryland and Federal Reserve Bank of New York. The views expressed in this paper are those of the individual authors and do not necessarily reflect the position of the Federal Reserve Bank of New York or the Federal Reserve System. Address correspondences to Joseph Tracy, Federal Reserve Bank of NY, Research Department, 33 Liberty Street, New York, NY 10045.
email: Joseph.Tracy@ny.frb.org.

Labor disputes are an intriguing feature of the landscape of industrialized economies. Economists have had a long-standing interest in formulating a framework for understanding and analyzing labor disputes. The development of noncooperative bargaining theory provided the tools for a theory of collective bargaining and labor disputes. A general aim of this theoretical development is to inform policy makers of the efficiency and equity effects associated with different labor laws and institutions that govern and shape the collective bargaining process. While this new literature is still evolving, it can already offer many insights into the interplay between policy and the bargaining process. In this chapter, we will provide a sketch of this new collective bargaining theory and illustrate its ability to aid in policy analysis. We will also relate the predictions of the model to existing empirical findings in the literature.¹

The collective bargaining process is complex and any model must of necessity involve many simplifying assumptions. Labor disputes likely arise from a wide range of causes. No model will adequately capture all of these forces. The aim of the researcher is to try and identify some of the central factors that shape the typical bargaining situation. Hicks (1932) concluded that strikes were largely the result of faulty negotiations. Ross (1948) stressed the divergent aims of the union leaders and the rank-and-file union members. Ashenfelter and Johnson (1969) formalized Ross' political model of unions into a theory of strikes.

We will focus in this chapter on another strand in the literature that stresses the role of private information. Our starting point is the assumption that the typical labor dispute arises from private information about some aspect that is critical to reaching an agreement, such as the firm's willingness to pay. Negotiators have an incentive to misrepresent their private information during the bargaining, and so some credible means of communicating this private information is necessary. Costly labor disputes serve this function. A firm with a high willingness to pay prefers to settle early at a high wage without a dispute, while a firm with a low willingness to pay prefers to endure a labor dispute in order to settle at a low wage. Labor disputes act as a signaling device for the firm's private information regarding its willingness to pay.

It is important to capture key institutional and legal features into the bargaining model. A common feature of collective bargaining in many countries is the ability of the union to continue to negotiate under an expired labor agreement. So long as the union does not call a strike, the

¹ There are a number of surveys of collective bargaining and strikes that provide useful background. See for example Kennan (1986), Kennan and Wilson (1989, 1990, 1993) and Card (1990a).

terms and conditions of the expired agreement continue to apply during the extended period of negotiations. As we discuss below, the union has a variety of ways of putting pressure on the firm during a contract extension. This implies that unions have an important decision to make as the expiration of the agreement approaches. If a settlement cannot be reached, the union must decide between calling a strike or extending the contract – which we will call a *holdout*. Our view is that unions often find it advantageous to put pressure on the firm while continuing to work. Early research using noncooperative bargaining models failed to recognize the union's threat choice and assumed that the union's only recourse was to call a strike (see Fudenberg *et al.*, 1985, Hart, 1989, and Kennan and Wilson, 1989). The model's ability to explain the data is greatly improved by incorporating holdouts into the bargaining process.

In the next section of the paper we will sketch out a basic bargaining model that includes the unions threat choice between strikes and holdouts. To start, we will assume that the threat payoffs are constant over time. Also, we will not incorporate any decision by the firm to replace striking workers. We will then extend the model to allow the threat payoffs to vary over the course of the dispute, and to allow the firm to decide whether to use temporary replacements in the event of a strike. We will then use the model and extensions to analyze several labor policies and generate predictions as to their impact on dispute incidence, dispute duration and the settlement wage. We will compare these predictions to empirical evidence in the literature, and illustrate how these empirical estimates can be used to analyze the efficiency and equity ramifications of labor policy.

II. A Model of Wage Bargaining

A union and a firm are bargaining over the wage to be paid during a contract of duration T . We focus on a single contract negotiation.² We simplify the model by assuming that only the firm has any private information. The union's reservation wage – what the workers will receive during a strike – is common knowledge. If a striking worker can secure alternative employment during the strike, then the reservation wage is the nonunion wage. Otherwise, the reservation wage is determined by the striking workers' access to unemployment insurance, welfare and/or

² We also ignore the timing of this negotiation relative to other related contract negotiations. The model applies most directly to bargaining in the private sector.

other union strike benefits. Let v be the firm's value of the current union labor force working under a contract of duration T . It is common knowledge that v is drawn from the distribution F with positive density f on an interval of support $[l, h]$. However, at the outset of the negotiations only the firm knows the realized value of v .

Negotiations begin with the union selecting a threat $\theta \in \{H, S\}$, where H indicates the holdout threat and S indicates the strike threat. The union's threat choice remains in effect until a settlement is reached. In the threat θ , the payoff to the union is x_θ and the payoff to the firm is $y_\theta(v) = a_\theta v - b_\theta$, where $a_\theta \in [0, 1)$ and $b_\theta \geq 0$. The term $1 - a_\theta$ captures the dispute cost in that threat. Define $c_\theta = (b_\theta - x_\theta) / (1 - a_\theta)$ to be the relative payment difference during the threat θ . Since the total payoff in agreement is v and the total payoff during the threat θ is $a_\theta v - b_\theta + x_\theta$, the "pie" that the union and firm are bargaining over (the difference between the agreement and the threat payoffs) is $(1 - a_\theta)v + b_\theta - x_\theta = (1 - a_\theta)(v + c_\theta)$. We assume that the pie is positive for all $v \in [l, h]$, which implies that $c_\theta > -l$.

For now, we will not be more specific about the strike treat ($x_S, a_S v - b_S$). However, the holdout threat can be further specified. Let w^0 denote the wage under the expired labor agreement. Since during a holdout the terms and conditions of the previous labor agreement remain in force, the workers continue to be paid w^0 , so $x_H = b_H = w^0$ and $c_H = 0$. We assume there is some inefficiency during a holdout, $a_H < 1$. There are several motivations for this inefficiency. First, we will show that during a holdout the workers have an incentive to slow down or "work to rule"; that is, to work exactly according to the stipulations of the expired agreement and no more. If firms rely on 'cooperation' of the workforce for efficient production, this may be an important source of inefficiency. Second, customers and input suppliers may be reluctant to negotiate new contracts with the firm until it has a new labor agreement in effect. Third, an issue in the current negotiations may be changes to the work rules which that will increase efficiency. These changes cannot be implemented until an agreement with the union is reached which stipulates the modified work rules. The precise degree of inefficiency during a holdout will turn out not to be important, so long as there is some inefficiency in this threat.

The outcome of this bargaining process between the union and the firm denoted by $\langle t, w, \theta \rangle$ consists of the time of the settlement t where $t \in [0, T]$, the wage settlement w and the threat selected by the union θ . The union and firm payoffs are calculated as the sum of the threat

payoffs and the agreement payoffs, weighted by the fraction of time spent in each outcome. We illustrate the union and firm payoffs in Figure 1 where we assume that both parties are risk neutral.

Figure 1 – Payoffs from Bargaining Outcome $\langle t, w, \theta \rangle$

Payoffs During Threat θ		Payoffs After Settlement	
Loss: $(1 - a_\theta)(v + c_\theta)$		Firm: $v - w$	
Firm: $a_\theta v - b_\theta$		Union: w	
Union: x_θ			
0	t		T
Old Agreement Expiration	Time of Settlement		New Agreement Expiration

Define

$$D(t) = \frac{1 - e^{-rt}}{1 - e^{-rT}}$$

to be the discounted fraction of time spent in dispute if an agreement occurs at time t . Then, given the bargaining outcome $\langle t, w, \theta \rangle$, the union's payoff is

$$U(t, w, \theta) = x_\theta D(t) + w [1 - D(t)]$$

and the firm's payoff is

$$V(t, w, \theta) = y_\theta D(t) + (v - w)[1 - D(t)].$$

The bargaining sequence is as follows. Following the union's threat choice the union and the firm alternate making wage offers, with the union assumed to make the initial offer. After a wage offer is made by one side, the other side has two options: (1) make a counteroffer, in which

case the bargaining continues, or (2) accept the current offer, in which case the bargaining ends and labor is supplied at the offered wage for the remainder of the contract period. As in Admati and Perry (1987), a bargainer can delay responding to an offer. This assumption leads to the signaling equilibrium in which the firm signals its value through its willingness to delay the agreement. For simplicity, we assume that the minimum time between offers is arbitrarily small.

The equilibrium of this bargaining game takes a simple form. If the wage under the expired labor agreement, w^0 , is sufficiently low (that is, below some indifference level \tilde{w}) the union decides to select the strike threat; otherwise ($w^0 \geq \tilde{w}$) the union selects the holdout threat. The indifference level wage, \tilde{w} , depends on r , T , F and the strike and holdout threat payoffs $[(x_\theta, y_\theta)$ for $\theta = \{S, H\}]$. A second indifference level, $m \in (l, h)$, determines the firm's response to the union's initial wage offer. If the firm's valuation is higher than this indifference level, $v > m$, the firm accepts the union's initial wage offer and an immediate settlement takes place without a dispute. Otherwise, the firm rejects the union's initial wage offer and a labor dispute begins. Whether the dispute is a strike or a holdout depends on the union's prior threat choice.

The signaling equilibrium can be characterized by three propositions from Cramton and Tracy (1992).³

Proposition 1. *Let $\theta = (x_\theta, y_\theta)$ be the threat chosen by the union. In the limit as the time between offers goes to zero, there is a perfect Bayesian equilibrium with the following form:*

- *The union makes an immediate offer of $w_\theta(m) = x_\theta + \frac{1}{2}(1 - a_\theta)(m + c_\theta)$, where $m(c_\theta) \in (l, h)$ maximizes*

$$(M) \quad (m + c_\theta)(1 - F(m)) + \int_l^m \frac{(v + c_\theta)^2}{m + c_\theta} dF(v).$$

³ Readers are referred to this article for further details and proofs.

- The firm accepts the offer if $v \geq m$. Otherwise, if $v < m$ the firm waits until $(m - v)/(m + c_\theta)$ of the contract period has passed before offering $w_\theta(v) = x_\theta + \frac{1}{2}(1 - a_\theta)(v + c_\theta)$, which is accepted by the union.
- The union's expected payoff from the threat θ is U_θ , the firm's expected payoff is V_θ , and the expected loss is L_θ , where

$$\begin{aligned}
 U_\theta &= x_\theta + (1 - a_\theta)(m + c_\theta)(1 - F(m)) \\
 V_\theta &= a_\theta E(v) - b_\theta + (1 - a_\theta) \int_m^h (v + c_\theta) dF(v) \\
 L_\theta &= (1 - a_\theta) \left[c_\theta - (m + 2c_\theta)(1 - F(m)) + \int_l^m v dF(v) \right].
 \end{aligned}$$

There are several observations that can be made based on Proposition 1. First, all wage offers are Rubinstein (1982) full information wage offers. The wage offer consists of the union's payoff in the threat θ plus half of the bargaining rents based on the firm's profitability, v .⁴ Second, during a holdout the union has every incentive to impose as much inefficiency on the firm as possible. The wage increases linearly with the degree of inefficiency, $w_H(v) = w^0 + \frac{1}{2}(1 - a_H)v$. However, several factors constrain the union in its endeavor to impose inefficiencies during a holdout. The union must not take actions that will put it in violation of the expired labor agreement. This gives meaning to the expression work-to-rule. Some countries give the firm the option of locking out the union. This option also limits the extent of inefficiency during a holdout. Finally, note that Proposition 1 generates a negatively sloped wage concession function. As the labor dispute continues, the union becomes more pessimistic about the firm's profitability (lowers its expectation regarding v). Since we have assumed that the threat payoffs and dispute costs remain constant during the dispute, this implies that wage offers

⁴ The union's initial wage offer is the Rubinstein wage offer for the firm type, m , that is indifferent between accepting or rejecting. The bargaining rents are split equally between the firm and the unions since we have assumed for simplicity that both parties have the same discount rate.

will be a declining function of the dispute duration.⁵ This result is analogous to the familiar union concession function of Ashenfelter and Johnson (1969).

For a given threat θ , we can determine how the dispute incidence and duration respond to changes in the threat payoffs and to changes in the distribution of v . The following proposition says that dispute activity increases with uncertainty. In addition, dispute activity increases when the threat θ becomes more attractive to the union (i.e. c_θ falls).

Proposition 2. *Suppose that $m(c_\theta)$ uniquely maximizes (M) . Dispute incidence $F(m(c_\theta))$ and dispute duration $D(v, c_\theta) = (m(c_\theta) - v) / (m(c_\theta) + c_\theta)$ increases as c_θ decreases. Likewise, dispute incidence and duration increase with a linear, mean-preserving spread of the distribution of F .*

Dispute activity depends on the amount of uncertainty about the firm's private information. Dispute incidence always exceeds one-half, and converges to one-half in the limit as this uncertainty disappears.⁶ Recall that c_θ measures what the firm pays less what the union receives in the threat θ scaled by the dispute cost. Proposition 2 yields several testable predictions. For example, if a local union receives strike benefits throughout a strike from its national union (and the costs of the benefits are spread across the national membership), then this lowers c_S which should increase strike incidence and lengthen strike durations. Similarly, if workers on strike qualify for general welfare payments, this also lowers c_S and should increase strike incidence and duration.⁷

Our third proposition demonstrates that the union's threat decision depends critically on the current wage under the expired labor agreement.

Proposition 3. *If $w^0 < \tilde{w}$, the union selects the strike threat; if $w^0 \geq \tilde{w}$ the union selects the holdout threat, where*

⁵ We will have more to say about the slope of the wage concession function later in this chapter.

⁶ By dispute incidence we mean the likelihood that either a strike or a holdout takes place.

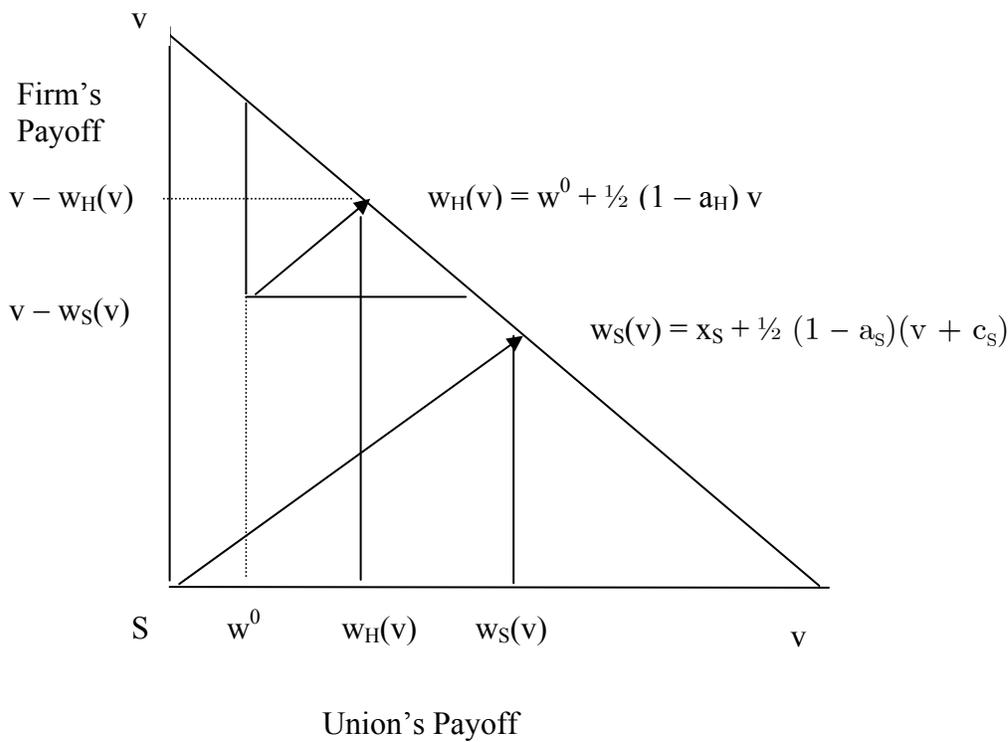
⁷ Until 1981, strikers in the United States were permitted to collect food stamps if they met the eligibility requirements. The General Accounting Office estimated that approximately \$37 million was paid out to striking households in 1980; for details, see Hutchens, Lipsky and Stern (1989).

$$\tilde{w} = x_s + (1 - a_s)(m(c_s) + c_s)[1 - F(m(c_s))] - (1 - a_H)m(0)[1 - F(m(0))]$$

and $m(c_\theta)$ maximizes **(M)**.

The intuition is that the union will select the strike threat if and only if the higher bargaining costs that are associated with a strike are more than made up for by a higher wage. If the current wage under the expired labor agreement is sufficiently high, this is not the case and the union prefers the holdout threat. This is illustrated in Figure 2.

Figure 2. The *Ex Post* Outcomes for Strike and Holdout Threats



In Figure 2, we have normalized the strike threat payoffs to the origin. If the strike threat (S) is selected, then the settlement occurs at the wage $w_S(v)$. Again, this settlement wage pays the union its threat value in the strike plus half of the inefficiency imposed during the strike. If the holdout threat (H) is selected, then the settlement occurs at the wage $w_H(v)$. The union will

select the strike threat only if the wage differential between the strike and holdout threats compensates the union for the additional dispute costs incurred in the strike threat. As the real wage under the expired labor agreement is increased, the wage differential between the strike and holdout threats becomes smaller, but the added dispute costs from the strike is unchanged. At some point, \tilde{w} , as w^0 continues to increase the union prefers to switch from the strike to the holdout threat.

The overall incidence of strikes depends not just on overall the incidence of disputes, but also on the fraction of disputes that involves a strike. As shown earlier, the level of dispute activity depends on the degree of uncertainty. The composition of disputes between strikes and holdouts depends on w^0 , the threat payoffs and the location of the distribution of v . Proposition 3 provides a key insight into strike activity. For example, the model predicts that the composition of disputes will shift toward strikes in the following situations: (i) after a period of uncompensated inflation which causes the real value of w^0 to decline, and (ii) when conditions in the local labor market are tight which raise the worker's reservation wages during the strike threat. Furthermore, if the level of uncertainty is relatively constant over time, then observed swings in strike incidence over time will primarily reflect shifts in the composition of disputes between strikes and holdouts.

An additional implication of Proposition 3 is that the structure of compensation in a labor agreement can affect the propensity for a strike to take place when the agreement comes up for a renewal. Whether the agreement incorporates a cost-of-living (COLA) clause and the degree of indexation built into the COLA will affect the union's holdout threat when the agreement expires. A prediction is that following a period of unexpected inflation renewals of nonindexed contracts will more likely involve a strike.⁸ Similarly, in the U.S. during the concession bargaining of the mid-1980s many firms successfully negotiated lump-sum wage payments in lieu of increases in the base wage (or in return for smaller base wage increases). This switch in pay structure implies a lower real wage at the end of the agreement, and consequently a stronger incentive for the union to select the strike threat when the contract is up for renewal.

⁸ If the inflation were anticipated, a nonindexed contract could factor the expected inflation into the deferred increases. In this case, we would not expect to see any differences in the likelihood of a strike between indexed and nonindexed contracts.

III. Extensions to the Wage Bargaining Model

The basic model provides several important insights into how the structure of the labor agreement and the prevailing economic conditions at the time of the labor negotiations affect the bargaining outcomes. However, to do a careful study of how labor policies affect the bargaining process, it will be helpful to generalize the model in two dimensions. The first generalization is to allow the threat payoffs to vary during the dispute. The second generalization is to incorporate the firm's decision on whether to use replacement workers in the event of a strike.

Time-varying Threat Payoffs

In labor contract negotiations, threat payoffs tend to change over time as inventories or strike funds are depleted, as strikers find temporary jobs, as public assistance to strikers is triggered by passing waiting periods for benefits and as replacement workers are hired and trained. With the exception of Hart (1989), most wage bargaining models assume that the threat payoffs to the union and the firm are constant during a dispute. This is a restrictive assumption when it comes to analyzing many labor policies which act to shift the threat payoffs during specific times in a dispute. Understanding how policy induced shifts in the threat payoffs affect the pattern of wage settlements and the level of dispute activity is important in evaluating the efficacy of these policies. In addition, this generalization will also sharpen our insights into how actions undertaken by the union and the firm such as inventory accumulation affect dispute activity.

To simplify, we consider a threat that consists of two phases indexed by $\theta \in \{1,2\}$, a short-run phase (1) and a long-run phase (2) with differing payoffs in each phase. Assume that the first phase of the threat lasts until time $\tau \in [0, T]$, after which the threat shifts to the second phase for the remainder of the contract period. In some cases, the switch time τ may be determined by the firm or the union, and in other cases it may be prespecified by a labor policy. Let the discounted fraction of the pie *remaining* at time t be defined as

$$d(t) = \frac{e^{-rt} - e^{-rT}}{1 - e^{-rT}}$$

and let $\delta = d(\tau)$. The payoffs to the union and the firm in each phase of the threat are illustrated in figure 3.

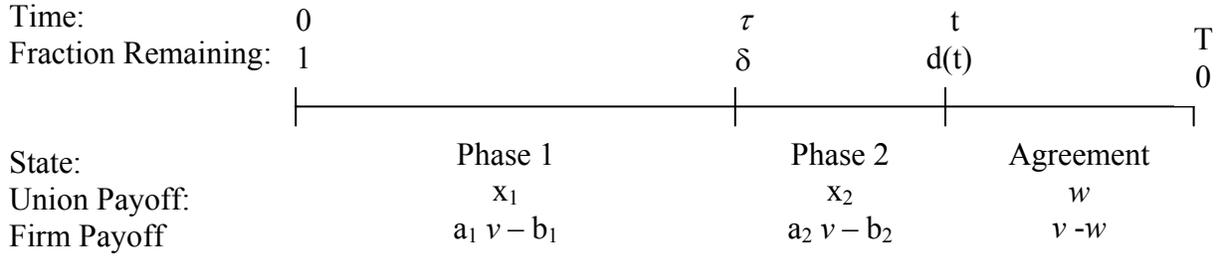


Figure 3. Two-phase threat payoffs.

A bargaining outcome, now denoted as $\langle t, w, \tau \rangle$, specifies the time of the agreement $t \in [0, T]$, the new contract wage w and the transition time between the short and long-run threats $\tau \in [0, T]$. The union's payoff is

$$U(t, w, \tau) = \begin{cases} (1-d(t))x_1 + d(t)w & \text{if } t \leq \tau \\ (1-\delta)x_1 + (\delta-d(t))x_2 + d(t)w & \text{if } t > \tau \end{cases}$$

and the firm's payoff is

$$V(t, w, \tau) = \begin{cases} (1-d(t))(a_1 v - b_1) + d(t)(v - w) & \text{if } t \leq \tau \\ (1-\delta)(a_1 v - b_1) + (\delta-d(t))(a_2 v - b_2) + d(t)(v - w) & \text{if } t > \tau. \end{cases}$$

The first step in the analysis is to generalize the Rubinstein (1982) full information wage offer in the presence of a two-phase threat. This generalization is given in Proposition 4 which is from Cramton and Tracy (1994b).

Proposition 4. *Suppose the value v is common knowledge, and the bargaining has reached time $t < \tau$ without a settlement. In the limit as the time between offers goes to zero, the full information wage at time t with a firm v is*

$$w(t, v) = (1-\gamma(t))w_1(v) + \gamma(t)w_2(v)$$

where

$$\gamma(t) = \frac{e^{-r\tau} - e^{-rT}}{e^{-rt} - e^{-rT}}, \quad w_\theta(v) = x_\theta + \frac{1}{2}(1-a_\theta)(v + c_\theta), \quad \theta \in \{1, 2\}.$$

The full-information wage from a sequence of threats is a weighted average of the full-information wages for each threat, where the weights are the fraction of the remaining discounted time spent in each threat, assuming that a settlement is never reached.

An important implication of Proposition 4 is that the settlement wage depends not only on the threat prevailing at the time of the settlement, but also on the unobserved threats that would have taken effect if the bargaining had continued. In our simple case of a two-phase threat, the settlement wage for an agreement that takes place during the short-run threat phase will be a weighted average of both the short-run threat wage and the long-run threat wage. If the time remaining until the shift in the threat is short relative to the remaining contract length, then the settlement wage is largely determined by the unobserved long-run threat. The following numerical example will illustrate this point. Suppose that the threat shift takes place after the first 30 days of the dispute (i.e. the firm's inventory runs out), the contract duration is 3 years and the discount rate is 10 per cent. Then at the outset of the bargaining, the long-run threat is given 97 per cent of the weight in determining the settlement wage. In this example, even if the dispute never reaches the long-run threat phase, the settlement wage is largely based on the long-run threat. So, the threat not reached can have an important influence on the wage settlement. Empirical wage specifications should take this feature of the model into consideration.

What is left to consider in a two-phase threat environment is the union's initial wage offer and the firm's response. Consider first the firm's response to the union's initial wage offer. The firm's optimal behavior is described in Proposition 5.

Proposition 5. *Suppose the union makes an initial wage offer of $w(m) = (1 - \delta)w_1(m) + \delta w_2(m)$, where $\delta = (e^{-r\tau} - e^{-rT}) / (1 - e^{-rT})$. In the subgame that follows, in the limit as the time between offers goes to zero, there is a perfect Bayesian equilibrium with the following form. If $v \geq m$, the firm immediately accepts the offer. If $v_\tau \leq v < m$, where*

$$v_\tau = m - (1 - \delta)(m + c_1) / \alpha,$$

and

$$\alpha = 1 + \delta \frac{a_1 - a_2}{1 - a_1},$$

then the second phase is not reached; the firm waits until

$$D(v) = \alpha \frac{m - v}{m + c_1}$$

of the contract has expired before offering

$$w(v) = (1 - \gamma(v))w_1(v) + \gamma(v)w_2(v),$$

where $\gamma(v) = \delta / (1 - D(v))$. If $v < v_\tau$, the second phase is reached; the firm waits until

$$D(v) = 1 - \delta \frac{v + c_2}{v_\tau + c_2}$$

of the contract has passed before making the offer $w_2(v)$. The firm's counteroffer is accepted immediately by the union.

To see how changes in the threat payoffs over time can affect dispute durations, consider first the case where the short-run dispute costs are low relative to the long-run dispute costs ($a_1 > a_2$). An example would be where the firm maintains its sales in the short-run by selling off inventory. In this case, $\alpha > 1$ and we see from Proposition 5 that dispute durations are longer than for the case of constant threat payoffs, $\alpha = 1$. This is the same as Hart's (1989) result based on a screening model where the strike costs are initially low but then increase sharply when a "crunch point" is reached. In contrast, when the short-run dispute costs are high relative to the long-run dispute costs, $\alpha < 1$ and dispute durations are shortened.

A prediction from the earlier constant threat payoff model was that settlement wages should decline with the duration of the dispute. As we will discuss later, some researchers have taken this as a central testable implication of noncooperative bargaining models of strikes. However, a downward sloping wage/strike duration concession function is not a robust prediction from these models. In the case of constant threat payoffs, whether wages rise or fall

with the dispute duration depends on where the private information exists. We have assumed that the firm had private information about its profitability and that the union's reservation wage was common knowledge. If we reverse the source of the private information, we also reverse the sign on the slope of the wage concession function. In addition, if we allow the threat payoffs to vary over the dispute, we can have an upward sloping concession function even if we maintain that only the firm has private information over its profitability. Consider the case where the short-run settlement wage is lower than the long-run settlement wage, $w_1(v) < w_2(v)$. As the dispute continues, the settlement wage is increasingly determined by the wage under the long-run threat. As a result, during the short-run threat phase the settlement wage increases with the dispute duration. These examples suggest that the slope of an estimated wage concession function is not a reliable test of noncooperative bargaining models of labor disputes.⁹

Fixed Costs and the Firm's Replacement Decision

While we could use the model with time-varying threat payoffs to speculate on the likely impact of a policy change regarding the firm's ability to use replacement workers, a fuller treatment of the firm's replacement decision is warranted. As a first step toward incorporating the firm's replacement decision into the model, we consider adding fixed costs to the threats. Hiring replacements involves a fixed cost by the firm, so it is important to know how fixed costs affect the bargaining process.

For simplicity, consider the model with constant threat payoffs. Now assume that there are fixed costs of initiating a dispute. Let $k_{f\theta}$ and $k_{u\theta}$ be the fixed costs for the firm and the union if threat θ is initiated, and let $k_\theta = k_{f\theta} + k_{u\theta}$ be the joint fixed cost. To make these fixed costs easily comparable to the flow payoffs (e.g. v and w), each k should be thought of as the flow k over the contract period that is equivalent to the one-time fixed cost K . These fixed costs can arise from costs of shifting production, a loss of goodwill during a dispute or the costs of looking for temporary employment. An immediate settlement avoids the fixed cost k_θ .

The presence of fixed costs alters the union's initial wage offer and the firm's decision whether to accept this wage offer. However, any counteroffer by the firm is independent of the

⁹ We will discuss in a later section some econometric issues that must also be taken into account when attempting to estimate the slope of the wage concession function.

fixed costs, since after the dispute has started the fixed costs are already sunk. These changes to the model are summarized in Proposition 6 taken from Cramton and Tracy (1998).

Proposition 6. *Let $\theta = (x_\theta, y_\theta)$ be the threat chosen by the union, where $y_\theta = a_\theta v - b_\theta$ and $c_\theta = (b_\theta - x_\theta)/(1 - a_\theta)$. In the limit as the time between offers goes to zero, there is a perfect Bayesian equilibrium with the following form:*

- *The union makes an immediate offer of $k_{f\theta} + w_\theta(m)$, where*

$$w_\theta(m) = x_\theta + \frac{1}{2}(1 - a_\theta)(m + c_\theta) \text{ and } m \in [l, h) \text{ maximizes}$$

$$k_{f\theta} - k_\theta F(m) + \frac{1}{2}(1 - a_\theta) \left[(m + c_\theta)(1 - F(m)) + \int_l^m \frac{(v + c_\theta)^2}{m + c_\theta} dF(v) \right].$$

- *The firm accepts the offer if $v \geq m$. Otherwise, if $v < m$, the firm waits until $(m - v)/(m + c_\theta)$ of the contract period has passed before offering $w_\theta(v) = x_\theta + \frac{1}{2}(1 - a_\theta)(v + c_\theta)$, which is accepted immediately by the union.*
- *If $m > l$, the union's expected payoff from the threat θ is U_θ , the firm's expected payoff is V_θ and the expected loss is L_θ , where*

$$U_\theta = k_{f\theta} - k_\theta [F(m) + (m + c_\theta)f(m)] + x_\theta + (1 - a_\theta)(m + c_\theta)[1 - F(m)],$$

$$V_\theta = -k_{f\theta} - k_\theta(m + c_\theta)f(m) + a_\theta E(v) - b_\theta + (1 - a_\theta) \int_m^h (v + c_\theta) dF(v),$$

and

$$L_\theta = k_\theta F(m) + 2k_\theta(m + c_\theta)f(m) + (1 - a_\theta) \left[c_\theta - (m + 2c_\theta)(1 - F(m)) + \int_l^m v dF(v) \right].$$

Once a dispute is initiated, the firm's counter-offer to the union and the associated delay in making this offer are both unaffected by the fixed costs. The intuition on how the fixed costs affect the union's initial wage offer is as follows. The union's initial wage offer is designed to make the firm with profitability m indifferent from accepting the wage offer, w , or rejecting and immediately making its counter-offer, $w_\theta(m)$. This indifference relationship is given by $m - w = m - w_\theta(m) - k_{f\theta}$. Rearranging gives $w = k_{f\theta} + w_\theta(m)$. The union, then, raises its initial wage offer by the amount of the firm's fixed cost.

While the firm's fixed cost in the threat θ affects the union's initial wage offer, the joint fixed cost affects the probability that this offer is accepted by the firm. This is consistent with the

joint cost hypothesis (Kennan, 1980; Reder and Neuman, 1980), which states that dispute incidence and duration should be inversely related to the joint cost of the dispute. Even a small fixed cost can have a sizeable impact on the likelihood of a dispute. For example, a fixed cost of roughly 1 per cent of the expected size of the pie will reduce the predicted dispute rate by 50 percent.

The fixed cost k_θ can also affect the union's threat decision. Recall that the union selects the strike threat if $U_S > U_H$. From Proposition 6 we see that fixed costs will shift the composition of disputes since U_S depends on k_{fS} and k_{uS} . As the union's fixed cost of striking increases, U_S falls, implying a composition shift towards holdouts. An increase in the firm's fixed cost of striking typically increases U_S , implying a composition shift toward strikes. However, the observed strike incidence depends not only on the composition of disputes, but also on the likelihood that the firm accepts the union's initial wage offer. Regardless of which party bears the fixed cost, observed strike incidence declines with an increase in fixed costs due to the increased likelihood that the firm will accept the union's initial wage offer.

Now consider extending the model to incorporate the firm's decision to use replacement workers. Replacement can be thought of as a new threat $\theta = R$ with its own payoffs. If the union selects the strike threat and the firm finds the union's initial wage offer unacceptable, then the firm decides whether to hire temporary replacements at a fixed cost k_{fR} . The fixed cost k_{fR} represents the hiring and training costs associated with replacement workers. We assume that replacements have a productivity level that is some fraction $a_R < 1$ of the productivity of the union workers.¹⁰

To see how the firm's replacement decision affects the bargaining process, consider first the full information case where v and k_{fR} are known by both the firm and the union. The full-information settlement wages under the three threats, $w_H(v)$, $w_S(v)$ and $w_R(v)$ play a key role in determining the outcome of the bargaining. If the cost of replacement is prohibitively high or the replacement wage is higher than the strike wage, $w_R(v) > w_S(v)$, then the firm never hires

¹⁰ We only consider here the decision to hire temporary replacements that are employed by the firm only until a settlement with the union is reached. In addition, we assume that the firm's decision to use temporary replacements does not affect v once a settlement is reached.

replacements. The settlement wage in this case is the strike wage, $w_S(v)$, or the holdout wage, $w_H(v)$, whichever is higher.

Now suppose that the replacement wage is less than the strike wage, $w_R(v) < w_S(v)$. This wage differential gives the firm an incentive to consider hiring replacements. If the cost of replacement is not too high so that $w_S(v) - w_R(v) > k_{FR}$, then if the union selects the strike threat it will reduce its wage demand in order to prevent the firm from hiring replacements. The union offers the wage $w = k_{FR} + w_R(v) < w_S(v)$, which makes the firm indifferent between accepting or rejecting and hiring replacements. While we never see replacements hired in the full information situation, the firm's option to hire replacements can lead to lower union wage demands in the strike threat.

To observe firms hiring replacements, we need to add in some private information about the firm's fixed cost of replacing its workers. To keep things simple, suppose that the firm's replacement cost is either low or high, $k_{FR} \in \{L, H\}$, and let $p = \Pr(k_{FR} = L)$ be the probability of a low replacement cost. Assume that $w_S(v) - w_R(v) \geq H$, so that the union must be concerned about replacement from both types of firms. The bargaining outcome may take the form of a pooling or a separating equilibrium. We consider each in turn.

Pooling equilibrium – If L and H are close, the union will make an initial wage offer that both types of firms will accept. In this case, the union makes the wage offer $w_L = L + w_R(v)$. This offer makes the low replacement cost firm indifferent between accepting and rejecting, and is strictly preferred by the high replacement cost firm. As in the full information case, we do not observe any replacements taking place, but the firm's option to replace reduces the union's wage demand in the strike threat. Consequently, the strike threat is less attractive to the union as a result of the firm's ability to hire replacements.

Separating equilibrium – As the gap between L and H widens, it becomes increasingly expensive for the union to buy out the replacement option for both types of firms. At some point, the union switches to the separating equilibrium where the union makes the initial wage offer $w_H = H + w_R(v)$ that makes the high cost firm indifferent between accepting and rejecting. A low replacement cost firm will reject this offer and subsequently hire replacements. The probability of observing replacements under the strike threat is p , which again is the probability

that the firm has a low replacement cost. The union's expected wage is now

$pw_R(v) + (1-p)w_H = w_R + (1-p)H$. The union's expected settlement wage is declining in the probability that the firm has a low replacement cost. Since the union only selects the strike threat if the strike wage exceeds the holdout wage by a sufficient margin, the union is less likely to select the strike threat as the replacement probability increases.

To observe both replacements and delays in settlements, we need to have private information both about v and k_{JR} . As long as the firm's decision to replace conveys only information about its fixed cost of replacement and not about v , adding the replacement decision does not upset our earlier equilibrium. Since replacement is a sunk cost, the signaling of this cost does not play a strategic role in the bargaining. In the separating equilibrium, the fraction of strikes involving replacements depends on the probability of a low replacement cost. In making its threat choice between a strike and a holdout, the union must estimate this replacement probability. If this fraction is too high, the union will select the holdout threat.

IV. Policy Analysis

We are now in a position to use this bargaining model and its extensions to analyze a range of labor policy issues. We will focus on the predictions that the model makes for dispute incidence and duration, the composition of disputes between strikes and holdouts and wage settlements. In the next section, we will compare these predictions with empirical findings in the literature. We will then illustrate how these empirical estimates can be used to evaluate the efficiency and equity properties of labor policies.

The most contentious labor policy issue over the past couple of decades is the right of employers to replace striking workers. We will analyze the implications for temporary replacements, leaving a discussion of permanent replacements for another occasion.¹¹ The firm's replacement option is expected to result in lower wage settlements for two reasons. First, if the union selects the strike threat, the model predicts that it will lower its wage demand in order to induce high replacement cost firms not to initiate replacements. Second, the replacement option reduces the attractiveness of the strike threat to the union, thereby shifting the composition of

¹¹ The model we presented only deals with a single contract negotiation and as such is not suited to analyze permanent replacements. In the United States, the incidence of temporary replacements is slightly higher than the incidence of permanent replacements (see Cramton and Tracy, 1998, pp. 674).

disputes to holdouts. Wage settlements are lower under the holdout threat, so a shift towards holdouts lowers the expected wage. To the extent that it takes time for the firm to hire and train temporary replacements once a dispute has begun, the replacement option makes the long-run dispute cost lower than the short-run dispute cost. From our discussion of time-varying threats, this should lead to shorter dispute durations. A ban on temporary replacement, then, should increase strike incidence both through a shift in the composition of disputes and the elimination of a source of fixed costs of strikes. In addition, a replacement ban should lead to longer dispute durations as well as higher wage settlements.

A number of Canadian Provinces have at various times imposed mandatory cooling off periods after an agreement has expired and before a strike can begin. We can use our two phase threat model to analyze the impact of a cooling off period. A cooling off period is a time interval where the union only has available the holdout threat. If a union selects the strike threat, then during the first phase it is restricted to the holdout threat until the cooling off period has been satisfied. In the second phase, the union switches to the strike threat. In practice, these cooling off periods tend to be short (typically less than three weeks' duration). As a result, Proposition 4 indicates that cooling off periods will leave wage settlements in the strike threat largely unaffected for those disputes that settle during the cooling off period.¹² Since with cooling off periods the dispute cost is higher in the second phase, Proposition 5 indicates that the settlement rate during the cooling off period will be lower than what would have occurred over this same time interval had the union selected the holdout threat. In this sense, cooling off periods have a "chilling" effect on the negotiations. A short cooling off period should not significantly alter the relative attractiveness of the strike threat, so this policy will not affect the composition of disputes. The prediction, then, is that short cooling off periods will have no impact on strike activity or wage settlements.

Unemployment benefits can also affect collective bargaining outcomes to the extent that striking workers qualify for these benefits. In the United States, eligibility for unemployment insurance (UI) benefits is determined at the state level. There are two common provisions that apply to strikes.¹³ The first is the *innocent bystander* provision that allows workers displaced by a strike and who have no direct connection or financial interest in the strike to collect UI

¹² From Proposition 4 it is also clear that wage settlements for observed strikes are invariant to the cooling off period.

benefits. The second is the *stoppage of work* provision that allows striking workers to collect UI benefits if the firm maintains its operations at or near normal levels during the course of the strike. In addition, New York (NY) and Rhode Island (RI) allow striking workers to file for UI benefits after a waiting period has passed.¹⁴

The impact of UI benefits on contract negotiations depends on which provision applies and how these benefits are financed. We illustrate these points by considering two of the three types of US statutes. The clearest predictions are for an innocent bystander provision. This provision provides UI benefits to workers at the firm who are *not* involved in the strike, so x_S is unchanged. If the firm is partially or fully charged for these benefits (this depends on the degree of experience rating in the state's UI system), then b_S increases which results in an increase in c_S . Conditional on the strike threat being selected by the union, Proposition 2 indicates that strike incidence and duration should decrease. However, the increase in c_S also raises the expected wage in the strike threat making this threat relatively more attractive to the union. Therefore, an increase in b_S generates two offsetting effects: a composition shift towards the strike threat (which acts to raise the observed strike incidence), and a higher likelihood that the union's initial wage offer will be accepted conditional on the union selecting the strike threat (which acts to lower the observed strike incidence). Cramton and Tracy (1994) find in their benchmark model that the net effect of an increase in b_S is a higher observed strike incidence. The magnitude of the effect of an innocent bystander provision on strike incidence should depend on the degree of experience rating in the UI system and the size of the UI benefits. In the extreme, if the firm is not charged for any of the UI benefits, b_S and c_S are unaffected and there should be no impact of the UI benefits on strike activity.

A distinguishing feature of the NY and RI provisions is that strikers qualify only after a substantial waiting period has elapsed. The two-phase threat model is useful for analyzing the effects of these waiting periods. UI benefits increase the union's reservation wage and therefore the settlement wage during the period when strikers qualify to receive the benefits. From

¹³ See Hutchens, Lipsky and Stern (1989) for a detailed discussion.

¹⁴ These waiting periods are 8 weeks for New York and 7 weeks for Rhode Island. Rhode Island also has an innocent bystander and stoppage of work provision that if applicable qualify workers for UI benefits after one week.

Proposition 4 we know that this will also increase wage settlements in a strike during the waiting period. Therefore, even if most strikes settle before the qualifying period for UI benefits has been satisfied, access to UI benefits will lead to higher wage settlements.

The implications of the NY and RI UI policies for strike incidence and duration are more difficult to discern. These implications will depend on the degree of experience rating in their UI programs, and may depend on the structure of unionization at the firm as well as actions taken by the firm during a strike. If the UI benefits are not charged back to the firm, then access to UI benefits reduces c_s during the period when the workers qualify for the benefits. From Proposition 2, this should make strikes more frequent and last longer. However, if most or all of the benefits are charged back to firms, then the impact on strike activity depends on additional factors. If the bargaining unit covers the entire firm, then the union's gain from UI is offset by the firm's increased UI taxes leaving c_s unchanged. In this case, UI benefits will not affect observed levels of strike activity. In contrast, if the firm has a significant number of nonunion workers, then the impact of UI on c_s depends on the firm's actions during the strike. If the firm maintains operations during a strike, then only strikers receive UI benefits and $b_s = x_s$ which again leaves c_s unchanged. On the other hand, if the firm shuts down operations as a result of the strike thereby idling its entire labor force, then UI benefits will be paid out to more than just the striking workers. In this case, $b_s > x_s$ which implies an increase in c_s and a reduction in observed strike activity.

As discussed earlier, firms have an incentive to accumulate inventories ahead of the expiration of a labor agreement. However, this strategy is not feasible in many service industries where no physical good is exchanged between the firm and its customers. Examples include the airline and shipping industries. The inability to use inventories in these industries raises the dispute cost (e.g. lowers a_s) making the strike threat attractive to unions. If permissible, firms in these industries may have an incentive to arrange "mutual aid pacts" among themselves to offset this disadvantage. The objective of these pacts is to lower the dispute cost via transfers between

the member firms and the firm experiencing a strike. By raising a_s these pacts should reduce strike activity and wage settlements.¹⁵

Firms can attempt to offset the cost of a strike in a number of ways. As noted already, firms can use inventory to offset production losses from the strike. In addition, a firm can attempt to use its existing nonunion workers (in many cases these are managers) to maintain production during the strike. This strategy tends to be more successful in industries that use a high degree of automation such as the telephone industry. Unions have an incentive to limit the firm's ability to follow this strategy by both making access to the plants difficult and by engaging in secondary picketing. To the extent that the union can increase the dispute cost in the strike threat by using these tactics, they increase the pie that is being bargaining over and consequently the expected wage settlements. Not surprisingly, in many countries the labor law deals extensively with defining the boundaries of appropriate actions that the union can engage in during a strike.

Finally, other government policies not directly aimed at regulating labor negotiations can have an impact on labor disputes. For example, consider wage and price controls. The model predicts that binding wage controls will reduce strike incidence and shorten dispute durations. Suppose that a union selects the strike threat but that its optimal initial wage offer $w(m)$ exceeds the allowed increase under the wage controls. The union will have to reduce $w(m)$ to satisfy the wage controls which increases the likelihood that the firm immediately accepts the wage offer. So, conditional on the union selecting the strike threat, the incidence of strikes will be lower with binding wage controls. In addition, binding wage and price controls should make the union less likely to select the strike threat. Since the initial wage offer under the strike threat is higher than the initial wage offer under the holdout threat, wage controls will be more binding for the strike threat than for the holdout threat. As a result, these controls will narrow the wage differential between the strike and the holdout threats making the strike threat less attractive. As unions

¹⁵ Six major airlines in the U.S. organized a Mutual Aid Pact (MAP) in October of 1958. The original agreement stipulated that a member who suffered a strike would be compensated by the other members a percentage of their "windfall" revenues resulting from the strike. In March of 1962, the agreement was changed so that a struck member would be guaranteed at least 25 per cent of its normal operating expenses. In 1969, this percentage was raised to a sliding scale from 50 to 35 per cent depending on the length of the strike. The MAP ended in 1978 with the deregulation of the airline industry. Over its lifetime, MAP transferred more than \$611 million dollars among its members (see Unterberger and Koziara, 1980).

substitute toward the holdout threat, the incidence of strikes is further reduced. Binding wage and price controls should also shorten average dispute durations as a direct consequence of the fact that more firms accept the union's initial wage offer. With more firms agreeing to a new contract at the outset of the bargaining, it will on average take less time for the remaining firms to signal their profitability to the union.

V. Empirical Evidence on Collective Bargaining and Strikes

The central feature of noncooperative bargaining models of labor disputes is the role of private information that creates uncertainty surrounding key issues in the negotiations. Labor disputes act as a credible way to signal this information. Ideally, this should be the focus of empirical tests of private information models. However, almost by definition, this is the most difficult test to carry out in the data. While there is ample anecdotal evidence supporting this view of labor disputes, formal statistical tests are difficult to construct.¹⁶ Tracy (1986, 1987) uses the volatility of a firm's stock return as a proxy for the union's degree of uncertainty. The premise of this test is that investors and the union negotiators share the same degree of uncertainty about the firm's future profitability. In a sample of US contract negotiations between 1973 and 1977 he finds that strike incidence has a positive and significant relationship to the variance of the firm's overall stock return. Furthermore he finds that this result is driven by the variability in the firm's excess returns and not by the variability in the market returns.¹⁷ Higher variance in the firm's stock returns is also associated with longer strike durations. A within sample one standard deviation increase in the stock return variability leads to an increase in the conditional strike duration by 8 days (the sample average conditional strike duration is 50 days), and to an increase in the unconditional strike duration by 2.5 days (the sample average unconditional duration is 7.5

¹⁶ The *New York Times* (October 28, 1990) described the negotiations between New York City and its municipal workers. 'At the core of the stalled New York City municipal labor talks is the union leaders' growing distrust of the claims being made about the city's financial plight.' ... 'Barry Feinstein, president of the 12,000-member Local 237 of the International Brotherhood of Teamsters, who is negotiating with Mr. Hill, said: "Any number of times the former administration would bring us in and say: 'This time, it's really true.' We would say, 'What makes this night different from all others?' On every occasion we were right, and the city was wrong." Mr. Feinstein called the city budget director, Philip R. Michael, "more full of it than a Christmas turkey." Mr. Michael insists his estimates are sound.'

¹⁷ This distinction is significant since variability in the excess returns reflects firm/industry uncertainty whereas variability in the market return reflects general macro uncertainty.

days).¹⁸ Cramton and Tracy (1994a) examine a sample of U.S. labor negotiations from 1970 to 1989 and find a significant relationship between a firm's stock return variance and *dispute* incidence in the 1970-81 period, but that this relationship disappears in the 1982-89 period.¹⁹ Finally, Ohtake and Tracy (1994) report contrasting results for Japan using industry level bargaining data from 1970 to 1990. They find that the Japanese industry average dispute incidence is positively related to macro uncertainty, but negatively related to industry uncertainty.²⁰

The degree to which strike incidence and duration display a cyclical pattern has received the greatest attention in the empirical literature. Early papers – such as Rees (1952) – examined the time-series properties of aggregate strike frequency. More recent work has focused on individual contract data in order to study the cyclical behavior of strike incidence. These studies are more informative for testing bargaining models.²¹ Researchers differ in their approach to testing for cyclical patterns to strike activity. These differences range from the specific choice of proxy for the business cycle and whether the cyclical variable is measured at a national or a industry/regional level.

The data generally indicate that strike duration is counter-cyclical and suggest that strike incidence may be pro-cyclical. Harrison and Stewart (1989, 1994) using Canadian data provide the most comprehensive analysis of the cyclical behavior of strike activity. The advantages of their studies are that they have a long time series that increases the cyclical variation in the data and they examine the robustness of their finding across several proxies for the business cycle including the choice of macro series and the level of aggregation.²² They find that strike durations are counter-cyclic. This finding is robust to using as the cyclical proxy aggregate or industry industrial production, real GDP or the male unemployment rate. Strike incidence appears to be pro-cyclical for manufacturing contracts based on aggregate industrial production.

¹⁸ The conditional strike duration is the duration given that a strike takes place. The unconditional strike duration includes negotiations that do not go on strike and assigns them a duration of zero days.

¹⁹ One hypothesis for why this relationship disappears is that the threat of replacements increased after President Reagan fired the air traffic controllers in August of 1981.

²⁰ Specifically, the industry dispute incidence is positively related to the risk-adjusted market return variability and negatively related to the industry excess return variability.

²¹ The connection between aggregate strike frequency and aggregate strike incidence can be weak. As discussed in Harrison and Stewart (1989), not all strikes take place at contract expirations and there may be cyclical fluctuations in the number of expiring labor agreements.

²² Harrison and Stewart examine Canadian strike data from 1946 – 83, and contract data from 1964 – 88.

In contrast, strike incidence for non-manufacturing contract negotiations displays no cyclical effects. When the male unemployment rate is used as the cyclical proxy, there is weak evidence that strike incidence is pro-cyclic – that is, strikes are less likely to occur in slack labor markets.²³ Controlling for aggregate industrial production, the authors find no evidence that industry-specific industrial production shifts strike incidence. Similarly, controlling for the aggregate male unemployment rate, they find no evidence that the Provincial unemployment rate shifts strike incidence.

Several other studies have examined the cyclical behavior of strike activity using data with a relatively long period of coverage. Cramton, Gunderson and Tracy (1999) also examine Canadian negotiations for the period 1967 to 1993. They control for the cycle using the Provincial unemployment rate and do not find any significant relationship with either strike incidence or duration. Vroman (1989) analyzes US manufacturing negotiations from 1957 – 84. She finds that strikes are more frequent but of shorter duration in periods of low unemployment. Finally, Ingram, Metcalf and Wadsworth (1993) examine 6,000 British manufacturing negotiations in the 1980s and find strike incidence to increase with slackness in regional labor markets. The authors speculate that the 1980s may not be a representative decade in Britain to test for cyclical strike effects.²⁴

While the evidence on procyclic strike incidence is far from definitive, it is a challenge for bargaining models to be able to explain why in "good times" strike incidence might increase but strike durations decrease. In the traditional bargaining model where the union has only the strike threat, the incidence and duration of strikes move in tandem. However, by introducing the union's threat choice, the model can reconcile why strike incidence and duration may move in opposite directions. Assume that during an economic expansion that the distribution of firm profitability, $F(v)$, shifts to the right. In addition, assume that the opportunities for striking workers to find temporary employment also improve. Proposition 2 states that the improved profitability will reduce both dispute incidence and dispute duration. That is, whether the union selects the strike or the holdout threat, the likelihood of observing a dispute (that is, the firm rejects the union's initial wage offer) falls and conditional on a dispute taking place the expected

²³ For the subset of strikes where wages are the primary issue, they find stronger statistical evidence that strike incidence falls in slack labor markets.

²⁴ For example, when they include an index to control for legislative changes the coefficient on the unemployment rate is reduced by two-thirds, although it is still positive and statistically significant.

duration of the dispute is shorter. At the same time, the enhanced temporary employment opportunities improve the union's strike payoffs but leave its holdout payoffs unaffected (the wage under the expired contract is fixed in advance). During business expansions, then, the strike threat becomes relatively more attractive to the union than the holdout threat which shifts the composition of disputes towards strikes. As a result of the union's threat choice, there are two offsetting forces at work on the cyclical behavior of the strike incidence – the union is more likely to select the strike threat (composition effect) and the probability of a strike is lower given that the strike threat is selected. Cramton and Tracy (1994a) show in their benchmark model that the composition effect dominates. Incorporating the union's threat decision, then, makes it possible for bargaining models to explain how strike incidence and duration can display opposite cyclical patterns.

The two-threat bargaining model also predicts that strike incidence increases when there is uncompensated inflation over the prior contract. Uncompensated inflation lowers the real wage under the expired contract making the holdout threat less attractive. The increase in strike incidence is the result of the shift in the composition of strikes. A model with only a strike threat will not generate this prediction. The real wage under the expired contract plays no significant role in a model where a strike is the union's only threat.

Several studies have controlled for measures of uncompensated inflation in the prior contract. Vroman (1989) in her study of US manufacturing contract negotiations finds a positive and significant impact of uncompensated inflation and the likelihood of a strike. Cramton and Tracy (1994a) find the same result in their study of US negotiations from 1970-89. This same pattern also shows up in Canadian contract negotiations; see Gunderson, Kervin and Reid (1986, 1989) and Cramton, Gunderson and Tracy (1999).

There has been little research to date that investigates the extent to which transfer payments affect strike activity. Skeels and McGrath (1997), using US data, include a measure of union liquid assets per member into a model of strike incidence.²⁵ They find that unions with more liquid assets available to finance strike benefits have a significantly higher strike incidence. Hutchens, Lipsky and Stern (1989, 1992) provide the most comprehensive work on government transfer payments and collective bargaining. They examine the impact of innocent bystander,

²⁵ They use liquid assets per member as a proxy for strike benefits since information on U.S. union strike funds is not widely reported.

stoppage of work and the NY/RI provisions on aggregate state strike frequencies.²⁶ They also control for variations in the maximum UI benefit, and where possible the degree of experience rating in the state's UI program. Their results indicate that both innocent bystander and stoppage of work provisions are associated with a higher strike frequency. They also find evidence that strike frequencies are higher in NY and RI, although the magnitude is not precisely estimated.

Several researchers have investigated whether wage and price controls have any effect on labor disputes. McConnell (1990) finds for the U.S. that the first stage of the Nixon wage and price controls reduced the incidence of strikes by nearly 5 percentage points (relative to an overall sample strike incidence of 13.8 per cent). She finds that the second stage of the Nixon controls had a smaller negative effect on strike incidence, but that this impact is imprecisely estimated. McConnell reports that both Nixon control periods appear to be associated with shorter strikes, but neither duration effect is well determined. Gunderson, Kervin and Reid (1989) report that the Canadian anti-inflation board was associated with a 4.3 percentage point reduction in strike incidence (relative to an overall sample strike incidence of 16 per cent). Using a longer time-period, Cramton, Gunderson and Tracy (1999) find that the Canadian anti-inflation board was associated with a 5.7 percentage point reduction in strike incidence and a 24 per cent reduction in strike durations.

We turn now to the empirical evidence on the impact of labor policy on the collective bargaining process, beginning with prohibitions on the use of replacement workers. The difficulty in testing predictions for replacement bans is the paucity of policy variation in the data. While several attempts have been made in the U.S. to pass replacement bans, none have been signed into law. Researchers have turned to Canadian data where three provinces have passed replacement bans.²⁷ Cramton, Gunderson and Tracy (1999) examine Canadian manufacturing and non-manufacturing contract negotiations from 1967 to 1993 and find that, following the enactment of a replacement ban, strike incidence is 12 percentage points higher, strike durations are on average 24 days longer and wages settlements are 4.4 per cent higher.²⁸ Gunderson, Kervin and Reid (1989) examine Canadian data that span a period prior to the adoption of

²⁶ The connection to the theory would be tighter if the empirical work related to strike incidence rather than strike frequency.

²⁷ Quebec has banned replacements since 1978 while British Columbia and Ontario introduced bans in 1993.

replacement bans by British Columbia and Ontario. They find a positive but insignificant effect of Quebec's ban on strike incidence. Gunderson and Melino (1990) estimate that the impact of Quebec's replacement ban was to increase strike durations by 7 days. Finally, Budd (1996) examines Canadian manufacturing contracts from 1965 – 85 and discerns no significant strike incidence, duration or wage effects associated with this replacement ban.

The bargaining model indicates that the higher the union's perceived risk of replacement, the less likely it is to select the strike threat. Cramton and Tracy (1998) test this prediction using data on US contract negotiations from the 1980s. For each of the major strikes in this period, they determine if temporary or permanent replacements were hired. They estimate the probability that a firm hires replacements (conditional on the union selecting the strike threat) using a probit model that controls for the likely demographic characteristics of the union workers and local labor market conditions. Consistent with the model, they find that the higher the "cost" of replacement, the less likely the firm is to use replacements.²⁹ Using this estimated probit model, they predict the replacement risk for each negotiation in their sample. They include this estimated replacement risk as a control variable in a probit model of the union's threat decision. The data indicate that unions facing higher replacement risks are significantly less likely to have selected the strike threat conditional on a dispute taking place.

Cooling-off periods have been mandated at some point by each Canadian Province. Both the period of coverage and the length of the cooling-off period varies across Provinces.³⁰ Cramton, Gunderson and Tracy (1999) find no significant effect of cooling-off periods on strike incidence or duration. Inconsistent with the model's prediction, they also find that 10 days of a required cooling off period is associated with 4.2 per cent lower wage settlements. Gunderson, Kervin and Reid (1989) also document the lack of any effect of cooling-off periods on strike incidence.

We suggested earlier that the slope of the wage concession function is not a robust test of bargaining models. In addition, there are econometric complications involved in estimating the slope of the wage concession function. McConnell (1989) using US data finds that wages on

²⁸ The strike incidence estimate is not precisely measured, while the strike duration and wage estimates are significant at the 10 per cent level.

²⁹ For example, the replacement risk is declining in the average experience of union workers in that industry and with the tightness of the local labor market.

average decline by 3 per cent per 100 days of strikes. Card (1990b) using Canadian contract negotiations finds no evidence of a negative relationship between wages and strike durations. Both studies control for bargaining unit heterogeneity and both treat the strike variables (an indicator for a strike and a measure of the strike duration) as exogenous. Jimenez-Martin (1999) using Spanish data, and Cramton, Gunderson and Tracy (1999) using Canadian data, reject this exogeneity assumption. Each of these studies finds an insignificant slope to the concession function when estimated with ordinary least squares, and a significant negative slope when estimated with instrumental variables. Care must be taken in estimating the relationship between wages and strike durations, and as we have already discussed care must be taken in interpreting the estimated results.

A common trend across industrialized countries has been the decline in strike activity over the 1980s and 1990s from earlier decades. Aggregate measures of strike activity will overstate this decline to the extent that unionization rates have also been receding during this period. However, even measures of strike incidence based on contract level data indicate a significant decline.³¹ While changes in the economic environment may explain some of this decline, researchers also suggest an important role for labor policy. Cramton and Tracy (1998) argue that there is suggestive evidence that President Reagan's decision in August 1981 to fire the air traffic controllers after their union, Professional Air Traffic Controller's Organization (PATCO), called an illegal strike contributed to the decline in US strike activity. To the extent that President Reagan's action signalled a greater public acceptance for the use of replacement workers, unions have an incentive to shift from strikes to holdouts. More directly, Ingram, Metcalf and Wadsworth (1993) find that changes in British labor law in 1982 and 1984 help to explain the decline in strike incidence in British manufacturing in the 1980s.

VI. Measuring the Effect of Labor Policy on Welfare

In this section we illustrate how the estimated impact of labor policy on strike incidence, strike durations and wage settlements can be combined to study the equity and efficiency implications of labor policy. The methodology is taken from Currie and McConnell (1999). We need to

³⁰ Cooling off periods have varied from 2 days in British Columbia to 21 days in Prince Edward Island and Nova Scotia.

³¹ For example, in the U.S. strike incidence among large bargaining units (1,000 and more workers) fell by roughly 50% from 1970 to 1990. See Figure 1 in Cramton and Tracy (1998).

assume a joint strike cost. Cramton, Gunderson and Tracy (1999) make the conservative assumption that the joint strike costs can be estimated by the lost wages (which amount to \$100 per person-day in their sample).

Labor policy can affect the efficiency with which labor agreements are renegotiated. Policies that lower the expected negotiation costs help improve bargaining ‘efficiency’. We will proxy for these negotiation costs with the expected strike costs. As we have seen, labor policy can affect strike costs through its impact on the incidence and the duration of strikes. We summarize this by looking at the policy impact on the unconditional strike duration, which is calculated as the product of the probability of a strike and the expected conditional strike duration. The strike cost associated with the policy, then, is the estimated change in the unconditional strike duration due to the policy multiplied by the mean daily strike cost.

Labor policy can also have equity effects by shifting the division of the rents between the union and the firm. We capture this equity aspect of labor policy by examining the change in the wage gain due to a policy. The wage gain from a policy is measured as the change in the wage bill over the life of the contract with and without the policy. This understates the full impact of the policy to the extent that wage gains in the current contract are carried over into future contracts. In calculating the net gain from a policy, we assume that the strike cost is split equally between the union and the firm. As a result, the union’s net gain is the wage gain less half the strike cost, and the firm’s net loss is the wage gain plus half the strike cost.

Table 1 illustrates this methodology by summarizing the equity and efficiency effects of cooling-off periods and replacement worker bans. These calculations are taken from Cramton, Gunderson and Tracy (1999). Cooling off periods have no significant impact on the efficiency of contract negotiations. However, by restricting the union’s threat choice, cooling-off periods shift the distribution of rents towards firms. Each week of required cooling-off is estimated to cost unions 2.6 million Canadian dollars per contract negotiation. This may be a good example of unintended policy consequences. A typical argument for cooling-off periods is that they will lower expected bargaining costs without imposing any terms on the firm or the union – that is, the policy will improve efficiency without any significant equity effects.

Looking at Table 1, it is easy to see why replacement legislation has been such a divisive policy issue. A ban on replacements is estimated to increase the expected bargaining cost by almost 2 million Canadian dollars per contract. In addition, a replacement ban is estimated to

raise the expected wage gain by 3.9 million Canadian dollars per contract. This large wage gain more than offsets the higher expected strike costs for unions, implying a net gain for unions of 2.9 million Canadian dollars per contract. Firms are significantly worse off with a replacement ban, with an estimated net loss of 4.8 million Canadian dollars per contract.

VII. Alternative bargaining models

We have focused on one simple bargaining model in which the firm signals its profitability through its willingness to postpone agreement. The advantages of the signaling model are that its equilibrium is tractable and the model is easily extended to allow other features, such as the choice of threat, more complex threats that change over time, as well as other forms of uncertainty.

A criticism of the signaling model is that it imbeds a particular type of commitment: the union is unable to improve its offer while it is waiting for the firm's response. In contrast, screening models allow the union and firm to revise their offers on a particular schedule. Typically, one assumes stationary strategies, which imply that the terms of agreement critically depend on the time between offers (Gul and Sonnenschein, 1988). This is problematic for empirical work, since it suggests that the main features of interest—strike incidence, strike duration, and wages—depend critically on the parties' ability to commit to not revising offers. Indeed, disputes vanish in the limit as the time between offers goes to zero. The fact that the time between offers is unknown to the researcher poses a serious challenge in applying the theory (but see Kennan and Wilson, 1989 and 1993, for a rich analysis and discussion of the issues).

At the opposite extreme from screening models are war-of-attrition models. These models effectively make the assumption that offers cannot be revised. Compromise is not possible. Either the union's terms or the firm's terms are ultimately adopted. The party willing to postpone settlement the longest ultimately prevails.³² The conceptual difficulty with attrition models is that compromise is a fundamental feature of collective bargaining. Although some issues may be binary, the most important, the wage, is not. As a result, attrition models provide some insight in understanding disputes, but have limited applicability to labor negotiations. In contrast to the attrition models, the ability to compromise is the primary determinant of wages in both the screening and signaling models.

We have emphasized the importance of the union's threat choice, either strike or holdout, in understanding collective bargaining data. In the theory, holdouts serve as an alternative, less destructive and safer, way for the union to put pressure on the firm. Holdouts can serve another useful purpose, enabling a union to delay negotiations until it obtains information about related bargaining outcomes in its industry (Kuhn and Gu, 1998). Holdouts of this form especially are important in situations where related bargaining pairs are negotiating simultaneously. A limitation of the basic model outlined in this chapter is that it treats each negotiation independently. In practice, negotiations can be linked by ties across bargaining pairs on either the firm or union side. Learning from these earlier negotiations can influence current negotiations (Kuhn and Gu, 1999).³³

A second important link in virtually all wage negotiations is the link from one contract to the next within the same bargaining unit. Union and firm bargaining relationships tend to be long-lived relative to any given contract duration. The prior contract is important, since it sets the status quo from which subsequent contracts are negotiated. Modeling a sequence of contract negotiations is difficult at best. Kennan (2001) characterizes equilibrium behavior in a setting where the firm has persistent information about its willingness to pay. The equilibrium involves information cycles that depend on the success of aggressive seller demands.

VIII. Conclusion

Over the last twenty years there has been a tremendous development of both the theoretical and empirical analysis of collective bargaining. Our focus here has been on work that seeks to explain the frequent occurrence of disputes. Our view is that disputes largely are motivated from the presence of private information and the sharply conflicting interests of the union and the firm over the wage.

To make sense of the empirical findings, we have found it important to extend the basic bargaining models to include key features of collective bargaining. One such feature is the union's threat choice: whether to withhold labor services through a strike or instead to put pressure on the firm while continuing to work, which we call a holdout. In this theory, changes in strike incidence are primarily the result of changes in the composition of disputes. More strikes

³² See Card and Olson (1995) for an application of this model to US strike data from the 1880s.

³³ See Wijngaert (1994) for an analysis of holdouts in the Netherlands.

occur when the strike threat becomes more attractive to the union. For example, striking becomes more attractive when there has been a significant amount of uncompensated inflation during the previous contract term, since this decreases the worker's wage in the holdout threat.

The bargaining model can also be extended to address the impact of other features of collective bargaining:

- threat payoffs that change over time, for example, as inventories or strike funds run out;
- the firm's choice to hire replacement workers;
- links among related bargaining pairs;
- links across contracts within the same bargaining pair.

The appropriate model to bring to the data depends in large part on the questions being asked and the institutional features of the country and industry being studied. Our focus here has been on private sector negotiations primarily in the U.S. and Canada. Although we suspect that many of the results apply elsewhere, it likely will be necessary to include other institutional details in interpreting collective bargaining in other country settings.

References

- Admati, Anat R., and Motty Perry (1987), 'Strategic Delay in Bargaining,' Review of Economic Studies **54**, 345-364.
- Ashenfelter, Orley, and George E. Johnson (1969), 'Bargaining Theory, Trade Unions, and Industrial Strike Activity,' American Economic Review, **59**, 35-49.
- Budd, John (1996), 'Canadian Strike Replacement Legislation and Collective Bargaining: Lessons for the United States,' Industrial Relations, **35**, 245-260.
- Card, David (1990), 'Strikes and Bargaining: A Survey of the Recent Empirical Literature,' American Economic Review, **80**, 410-415.
- (1990) 'Strikes and Wages: A Test of an Asymmetric Information Model,' Quarterly Journal of Economics, **105**, 625-659.
- Cramton, Peter, and Joseph S. Tracy (1992), 'Strikes and Holdouts in Wage Bargaining: Theory and Data,' American Economic Review, **82**, 100-121.
- (1994), 'The Determinants of U.S. Labor Disputes,' Journal of Labor Economics, **12**, 180-209.
- (1994), 'Wage Bargaining with Time-Varying Threat Payoffs,' Journal of Labor Economics, **12**, 594-617.
- (1998), 'The Use of Replacement Workers in Union Contract Negotiations: The U.S. Experience, 1980-1989,' Journal of Labor Economics, **16**, 667-701.
- Cramton, Peter, Morley Gunderson, and Joseph Tracy (1999) 'The Effect of Collective Bargaining Legislation on Strikes and Wages,' Review of Economics and Statistics, **81**, 475-487.

- Currie, Janet, and Sheena McConnell (1991), 'Collective Bargaining in the Public Sector: The Effects of Legal Structure on Dispute Costs and Wages,' American Economic Review, **81**, 693-718.
- Durgan, J.W., and W.E.J. McCarthy (1974), 'The State Subsidy Theory of Strikes: An Examination of Statistical Data for the Period 1956-1970,' British Journal of Industrial Relations, **12**, 26-47.
- Fudenberg, Drew, David Levine, and Paul Ruud (1985), 'Strike Activity and Wage Settlements,' Working Paper. University of California at Berkeley.
- Gul, Faruk, and Hugo Sonnenschein (1988), 'On Delay in Bargaining with One-Sided Uncertainty,' Econometrica, **56**, 601-611.
- Gunderson, Morley, John Kervin, and Frank Reid (1989), 'The Effects of Labour Relations on Strike Incidence,' Canadian Journal of Economics, **22**, 779-794.
- , and Angelo Melino (1990), 'The Effects of Public Policy on Strike Duration,' Journal of Labor Economics, **8**, 295-316.
- Harrison, Alan, and Mark Stewart (1989), 'Cyclical Fluctuations in Strike Durations,' American Economic Review, **79**, 827-841.
- (1994), 'Is Strike Behavior Cyclical?' Journal of Labor Economics, **12:4**, 524-553.
- Hart, Oliver (1989), 'Bargaining and Strikes,' Quarterly Journal of Economics, **104**, 25-44.
- Hicks, John R. (1932), The Theory of Wages, London: Macmillan Press.
- Hutchens, Robert, David Lipsky, and Robert Stern (1989), Strikes and Subsidies: The Influence of Government Transfer Programs on Strike Activity, Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.
- (1992), 'Unemployment Insurance and Strikes,' Journal of Labor Research, **13**, 337-354.
- Jimenez-Martin, Sergi (1999), 'Controlling for Endogeneity of Strike Variables in the Estimation of Wage Settlement Equations,' Journal of Labor Economics, **17**, 583-606.
- Kennan, John (1980), 'Pareto Optimality and the Economics of Strike Duration,' Journal of Labor Research, **1**, 77-94.
- (1986), 'The Economics of Strikes,' in Orley C. Ashenfelter and Richard Layard (eds.), Handbook of Labor Economics, **2**, 1091-1137. Amsterdam, North Holland, 1986.
- (2001), 'Repeated Bargaining with Persistent Private Information,' Review of Economic Studies, **68**, 719-755.
- , and Robert Wilson (1989), 'Strategic Bargaining Models and Interpretation of Strike Data,' Journal of Applied Econometrics, **4**, S87-S130.
- (1990), 'Can Strategic Bargaining Models Explain Collective Bargaining Data?' American Economic Review, **80**, 405-409.
- (1993), 'Bargaining with Private Information,' Journal of Economic Literature, **31**, 45-104.
- Kuhn, Peter, and Wulong Gu (1998), 'A Theory of Holdouts in Wage Bargaining,' American Economic Review, **88**, 428-449.
- (1999), 'Learning in Sequential Wage Negotiations: Theory and Evidence,' Journal of Labor Economics, **17**, 109-140.
- McConnell, Sheena (1989), 'Strikes, Wages, and Private Information,' American Economic Review, **79**, 810-815.
- Ohtake, Fumio, and Joseph Tracy (1994), 'The Determinants of Labour Disputes in Japan: A Comparison with the U.S.' in Toshiaki Tachibanki (eds.), Labour Market and Economic Performance: Europe, Japan and the USA, London: Macmillan, pp. 349-372.

- Reder, Melvin, and George Neumann (1980), 'Conflict and Contract: The Case of Strikes' Journal of Political Economy, **88**, 867-886.
- Rees, Albert (1952), 'Industrial Conflict and Business Fluctuations.' Journal of Political Economy, **60**, 371-382.
- Ross, Arthur M. (1948), Trade Union Wage Policy, Berkeley: University of California Press.
- Rubinstein, Ariel (1982), 'Perfect Equilibrium in a Bargaining Model,' Econometrica, **50**, 97-109.
- Skeels, Jack W., and Paul McGrath (1997), 'The Effect of Union Financial Strength and Liquidity on Strike Propensities,' Journal of Economics, **23**, 59-71.
- Tracy, Joseph (1986), 'An Investigation into the Determinants of U.S. Strike Activity,' American Economic Review, **76**, 423-436.
- (1987), 'An Empirical Test of an Asymmetric Information Model,' Journal of Labor Economics, **5**, 149-173.
- Unterberger, S. Herbert, and Edward Koziara (1980), 'The Demise of Airline Strike Insurance,' Industrial and Labor Relations Review, **34**, 82-89.
- Vroman, Susan (1989), 'A Longitudinal Analysis of Strike Activity in U.S. Manufacturing: 1957-1984,' American Economic Review, **79**, 816-826.

Table 1. Estimates of the Impact of Labor Policy on Welfare

Policy variable	Per cent change in wage	Change in unconditional duration (days)	Million June 1993 Canadian dollars per contract			
			Strike cost	Wage gain	Union net gain	Firm net loss
Cooling-off period (weeks)	-2.9	-0.3	-0.04 (0.04)	-2.62** (0.97)	-2.60** (0.97)	2.64** (0.97)
Prohibit replacement workers	4.4	15.1	1.93** (1.01)	3.88** (1.66)	2.91* (1.74)	-4.84** (1.74)

Notes: See Cramton, Gunderson and Tracy (1999). Standard errors are given in parentheses and reflect sampling variation in the estimated marginal effects of a policy variable on wage settlements, strike incidence and strike duration. Overall impacts are based on a mean strike incidence 0.165, a mean strike duration (D) of 59 days, a mean contract duration (l) of 822 days, a mean real wage (w) of \$13.81 (June 1993 Canadian dollars), a mean bargaining unit size (n) of 1,531 workers, a mean person-days lost due to strikes (d) of 76 thousand days, a mean number of hours worked per week (h) of 36, and a mean loss per strike ($L = w \cdot d \cdot h / 5$) of \$7.50 million, equal to the lost wages. Let D_i be the change in the unconditional strike duration. Strike costs = $C_i = L \cdot D_i / D$; wage gain = $W_i \cdot n \cdot l \cdot h / 5$; union net gain = $G_i - C_i / 2$; firm net gain = $-G_i - C_i / 2$.

** statistically significant at the 0.05 level. * statistically significant at the 0.10 level.