NBER WORKING PAPER SERIES

A BIOLOGICAL MODEL OF UNIONS

Michael Kremer
Benjamin A. Olken

Working Paper 8257
http://www.nber.org/papers/w8257

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2001

We are very grateful to Marcos Chamon and Alexander Veytsman for excellent research assistance; to Abhijit Banerjee, Ricardo Caballero, John Dunlop, Susan Dynarski, Robert Gibbons, Lawrence Mishel, Michael Piore, Andrei Sarychev, Jean Tirole and Michael Wallerstein for helpful discussions; and to Richard Freeman for conversations which inspired the paper. None of them should be held responsible for the results. The views expressed herein are those of the authors and not necessarily those of the National Bureau of Economic Research.

A Biological Model of Unions
Michael Kremer and Benjamin A. Olken
NBER Working Paper No. 8257
April 2001
JEL No. J51, E24, B41

ABSTRACT

This paper applies principles from evolutionary biology to the study of unions. We show that unions which maximize the present discounted wages of current members will be displaced in evolutionary competition by unions with more moderate wage policies that allow their firms to live longer. This suggests that unions with constitutional incumbency advantages that allow leaders to moderate members' wage demands may have a selective advantage. The model also suggests that industries with high turnover of firms will have low unionization rates, and that there may be one equilibrium with high unionization and long-lived firms and another with low unionization and short-lived firms. These predictions seem broadly consistent with the data.

Michael Kremer
Harvard University
Department of Economics
Littauer Center 207
Cambridge, MA 02138
mkremer@fas.harvard.edu

Benjamin A. Olken
Harvard University
GSAS
Cambridge, MA 02138
bolken@fas.harvard.edu

We compare a model of unions based on principles from biology with an economic model in which unions maximize the present discounted value of rents obtained by members. Biological models suggest that selection pressure often works against organisms that are too harmful to their hosts. For example, a disease such as the Ebola virus, which kills its hosts in days, has little opportunity to spread from one host to another. In contrast, the viruses that cause the common cold are widespread. Mitochondria, which were probably originally parasites, evolved to become essential to their hosts and are now universal.

This paper applies this basic biological concept to the interaction between unions and firms. We argue that unions that demand the level of wages optimal for their members will be displaced in competition with more moderate unions. In our model, the dynamics of union coverage depend on both the rate at which unions spread to non-union firms and the rate at which unionized firms die. Greater rent extraction by a union can increase the spread rate of the union by making the union more attractive to workers in non-unionized firms. On the other hand, greater rent extraction by unions can also lead to increases in the death rates of unionized firms. The evolutionarily stable level of rent extraction therefore involves a tradeoff between attractiveness to workers and the effect on firm death rates.

In the model, reducing the level of rent extraction slightly from the level that maximizes the present discounted value of rents to union members causes a second-order reduction in members' welfare, and hence in the spread of unions, but a first-order reduction in the death rate of unionized firms. Selection pressure therefore favors unions with lower levels of rent extraction than would be optimal for workers.

If a union is controlled by its rank and file, however, its members will vote for the policies that maximize their welfare rather than the policies that would survive evolutionary competition. Selection pressure may therefore also favor unions with constitutional incumbency advantages that allow leaders to pursue more moderate wage demands than those preferred by the rank and file.

There is evidence for these implications. Most existing unions do, in fact, have constitutions that create strong advantages for incumbents. Furthermore, rank and file dissident movements are almost always more radical than union leadership, suggesting that the policies of unions tend to be more moderate than would be optimal for workers. In several cases in which incumbency advantages have been weakened due to plausibly exogenous factors, dissident movements have become powerful, wage demands have escalated, and industries have declined. The model also suggests that if multiple unions compete for the same workers within firms, as in several European countries, incumbency advantages will be weaker and unions will have to adopt more militant policies.

The model also suggests that exogenous increases in firm turnover will reduce unionization. If the typical unionized firm exits before dues from union members at that firm have financed unionization of another firm, then unionization rates will decline. In fact, industries with high firm turnover, such as restaurants, have lower unionization rates than industries with lower firm turnover, such as hotels. We demonstrate empirically that this is the case even after controlling for capital intensity, concentration, and establishment size, and argue that the relationship may help explain higher unionization rates in the public sector, as well as the relatively weaker incumbency advantages found in public sector unions.

Although unionized firms are more likely to exit, in general equilibrium high unionization can reduce firm turnover. Since unions extract more in absolute terms from newer, more productive firms than from unproductive firms near the exit margin, unions may deter entry by new, high-productivity firms without extracting so much from older low-productivity firms that they exit. (See Moene and Wallerstein [1993] and Caballero and Hammour [1998].) Combining this finding with this paper's result that higher firm turnover reduces unionization suggests that there may be multiple steady states, one with high unionization, low turnover of firms, and low productivity and another with low unionization, high turnover of firms, and high productivity. This may correspond to some differences between the U.S. and Europe.

This paper builds on earlier work. Dickens and Leonard [1985] and Freeman [1983] show that unions must continually organize new enterprises in order to offset the natural decline in membership due to turnover among firms. Freeman [1998] documents sudden spurts in unionization followed by gradual declines. He accounts for this in a model in which as unionization levels increase, it becomes first easier and then more difficult to unionize new firms. This means that there will be one steady-state level of unionization at zero, and one positive steady state. Hannan and Freeman

[1987, 1988] use a sociological model of organizational ecology to examine how birth and death rates of unions depend on the existing number of unions. This paper differs in explicitly examining the predator-prey population dynamics involving unions and firms and in deriving the implications for union politics. Dutta and Radner [1999] apply related evolutionary techniques to a different question, arguing that firms that retain more earnings than would be optimal for their shareholders will survive longer and eventually outnumber firms that retain the optimal amount. This paper differs in methodology from Dutta and Radner, however, by explicitly modelling the spread of unions within a population of firms and by considering competition among unions in determining which unions will survive.

Section 1 provides background on relevant U.S. collective bargaining institutions. Section 2 presents the model and solves for the steady-state level of unionization with a single union and an exogenously given level of rent extraction. Section 3 contrasts economic and evolutionary models of the determination of levels of rent extraction, and finds that the evolutionarily stable level of rent extraction is less than that which maximizes the welfare of the workers. Section 4 discusses the model's implications for union institutions. Section 5 empirically tests the model's prediction that unionization levels will be lower in industries with greater firm turnover. Section 6 discusses two extensions of the model, one showing that evolutionarily stable unions may devote more resources to organizing efforts than would be optimal for their members and one showing that there can be multiple equilibria in unionization levels and turnover among firms. Section 7 concludes by arguing that the welfare effects of unions, and of union moderation, are ambiguous under the model, and by discussing the applicability of this biological approach to other institutions.

1 Background on U.S. Collective Bargaining Institutions

Before introducing the model, it is useful to review a few features of U.S. collective bargaining institutions. Outside of construction, music, and a few other industries, most new firms begin life without unions. Under the Federal law covering most industries, if thirty percent of workers sign a petition calling for an election, a certification election supervised by the National Labor Relations

Board (NLRB) is held. A union is recognized if more than half the workers vote for it in such an election.

Support from existing unions plays an important role in unionizing new firms. Not only are workers more likely to support unions if they have friends or relatives who are union members, but hired union organizers, paid for through dues of existing union members, also play an important role. These paid organizers are often critical in obtaining the signatures required to have an election and in campaigning for union certification, because unlike activists within firms, paid organizers are not susceptible to threats from management. Workers at a plant are theoretically protected from retaliation for supporting a union, but penalties for dismissing union supporters are weak, and union activists are often dismissed. In fact, one in twenty workers who vote for a union in an organizing election are later found to have a valid claim for unfair dismissal by the NLRB [Weiler, 1984]. The percentage among union activists is likely to be even higher, making it dangerous for workers in a firm to openly campaign for a union in an NLRB election. In addition to making organizing activities hazardous for employees, firms also use legal tactics to delay unionization votes, such as challenging definitions of the bargaining unit and thus the set of workers who are eligible to vote in the NLRB election. Responding to these challenges requires lawyers and money, which existing unions can help provide.

Once a firm unionizes, workers can theoretically deunionize through a decertification election, or vote to change their affiliation from one union to another. In practice, however, decertifications are infrequent, and switching union affiliations rarely happens, given the organizing costs involved and the reluctance of unions to poach each others' territory. In fact, the AFL-CIO constitution explicitly prohibits member unions from attempting to organize a firm currently organized by a different AFL-CIO member union. When unions decline, it is therefore not primarily because of decertification elections, but rather because the firms covered by the union reduce employment or close down a unionized location altogether.

The model in this paper is designed to apply to those U.S. industries covered by the standard NLRB rules: new firms start as non-union; paid union organizers play an important role in unionizing new firms; and once employees at a firm vote in a particular union, the firm stays unionized for the remainder of its life.[1] The resulting dynamics of unionization levels bear a similarity to those under the Susceptible-Infected (SI) model of epidemiological dynamics (see Anderson and May [1991]). In that model, new potential hosts are born uninfected; the chance that they become infected increases with the number of hosts already infected; and once hosts are infected, they stay infected until they die. (As discussed in the conclusion, this comparison is purely positive, not normative.)

2 The Model with a Single Union

This section describes the basic model for the spread of a single union with an exogenously given level of rent extraction. Section 2.1 begins by outlining the entry, investment, and exit behavior of firms taking union behavior as given. Section 2.2 then describes how unions spread and characterizes the steady-state level of unionization.

2.1 Firms

We assume that firms produce one of a continuum of measure F possible products, and that there is a downward-sloping demand curve for each product.[2] Entry into a sector requires start-up costs, described below, but once these costs have been paid, output is linear in labor and requires no other inputs, i.e. $q(L) = \beta L$. Once there is a firm in a market, if a second firm were to enter, the two firms would engage in Bertrand competition and earn zero profits. Knowing this, only one firm enters each market, and the measure of the number of firms is equal to F. For simplicity, we will assume that all firms face identical production functions, and so behave identically.

In addition, there is a competitive, constant returns to scale home-production sector in which workers can earn some fixed effective wage, $\underline{w}$. We assume that there is a sufficient quantity of workers such that some are always employed in the home-production sector, i.e. $N > L^* F$, where N is the quantity of workers and L^* is the optimum quantity of workers each firm employs at wage

[1] As discussed above, in a few industries, such as construction, textiles, and music, institutions differ, and new firms often start out unionized. The model is not intended to apply to these industries.

[2] One could instead assume that the number of products in the world was growing at a constant rate, which would induce steady-states of constant growth rates, but that change does not affect the basic intuition of the model.

$\underline{w}$.

Given that each firm is a monopoly, each firm charges the profit maximizing price and earns pre-union profits denoted by π. By "pre-union profits," we mean the surplus of revenues over the wages paid in the absence of a union. (We assume that there is some demand for each product at a price above $\frac{\underline{w}}{\beta}$, so that each firm produces a positive amount, and that the profits are maximized at some finite price.[3]) If the firm is unionized, the union extracts a fixed proportion α of these profits. Later, we will endogenize α, but from the perspective of the firm, α is an exogenous parameter.

Suppose that firms are subject to large negative productivity shocks that cause them to exit with hazard rate δ, where δ depends in part on unobservable investment, I, such as avoiding negligence that could lead to lawsuits.[4] For now, we will assume that the negative shocks are so large that any firm receiving such a shock exits; Section 6.2 considers smaller negative shocks that reduce productivity but are not necessarily fatal to firms. We also assume that $\delta_I < 0$ and $\delta_{II} > 0$.

The optimal investment for a unionized firm depends on the share of profits it can keep if it stays alive. Given the discount rate, r, the firm chooses I to maximize its present discounted value given α:[5]

$$I(\alpha) = \underset{I}{\operatorname{argmax}} \frac{(1-\alpha)\pi - I}{r + \delta(I)}. \tag{1}$$

Investment is decreasing in rent extraction by unions, α, since

$$\frac{dI}{d\alpha} = \frac{\pi \delta_I}{\delta_{II}\left[(1-\alpha)\pi - I\right]} < 0. \tag{2}$$

It is therefore possible to write $\delta = \delta(I(\alpha))$, or more concisely, $\delta = \delta(\alpha)$, where $\delta_\alpha > 0$.

So far, we have said that there will be only one firm in each industry, but have not yet specified how a given capitalist gets to own that firm. We model the process by which a given capitalist

[3] For example, suppose that all consumers had an identical CES utility function equal to $U = \left(\int_0^F x_i^\rho di\right)^{\frac{1}{\rho}}$, where x_i represents demand for good i. As long as $\rho > 0$, so that the elasticity of substitution is greater than 1, all firms will charge a finite price.

[4] The hazard rate could also depend on observable investment, but since unions and firms can contract on the efficient level of observable investment, it would not vary with rent extraction, and hence we abstract from observable investment in this paper.

[5] Note that equation (1) assumes that the owner of the firm receives a continuation payoff of 0 in the event the firm dies. This is because if the firm dies, the owner will need to start a new firm, and as will be shown below, the *ex-ante* profits of starting a new firm will be 0.

obtains the monopoly on a particular product as an auction or, equivalently, as a lottery. This can be thought of either literally, such as a government auction for a cell-phone license, or as a metaphor for advertising, research and development, or other up-front expenditures that result in some probability of being successful in an industry, as is widespread among Internet firms today. Assuming that there is competition among a large number of risk-neutral capitalists, the cost of entering an industry will be equal to the expected value of owning a firm. The *ex ante* profits from opening a firm will therefore always be zero. Whenever a firm dies, an auction is held and a new firm enters. The number of firms therefore remains equal to F.

2.2 Steady-State Unionization Levels

Under the model, new firms are established without unions. Firms differ in how easy they are to unionize, depending on factors ranging from the layout of the factory floor to the personalities of managers. (In order to keep the model tractable, we consider a simple model in which firms, plants, and union bargaining units are coterminous.) Each firm is born with a certain difficulty of being organized, which we denote by c (for cost), and retains that same level of difficulty until it dies. For simplicity, we will assume that for newborn firms c is distributed uniformly on the interval [0,1].

In each unit of time, the union has an organizing budget that it uses to organize new firms. We assume that unions are credit constrained, so that the amount they can spend on organizing efforts depends on their current level of dues collection. The union's budget is therefore equal to BU, where B represents the amount that unionized workers in each firm contribute toward the overall union's organizing budget and U is the number of unionized firms. (We abstract from size differences among firms.)

The attractiveness of a particular union to workers depends on α, the proportion of the firm's total profits it extracts for the workers. The union's effective organizing budget is $A(\alpha)BU$, where $A(\alpha)$ indicates the union's attractiveness as a function of α. Workers recognize that firms will die off quickly if unions extract high levels of rents, so $A(\alpha)$ will not necessarily be monotonic in α.

The analysis in this section will focus on identifying steady states. The transition dynamics outside of the steady state are somewhat more complex, and are discussed in Appendix A.2. There are two criteria that must be satisfied in the steady state. First, in the steady state, the total number of unionized firms, denoted U, must remain constant. Next, note that when a firm dies, the firm that replaces it has a new difficulty of unionization c, distributed according to the initial Uniform[0,1] distribution. This leads to the second criteria for the steady state, that the distribution of organizing difficulties of union and non-union firms must also remain constant.

To identify the steady state, in this section we first consider the case in which there is only one union, with an exogenously given level of α. (Section 3 endogenizes α.) We assume that the union can observe the difficulty of organizing a firm before it starts an organizing effort. Therefore, the union will target those firms that are the easiest to organize first.[6] Suppose that at a given moment all firms with organizing difficulty below some cutoff point p are unionized and all firms with difficulty above p are non-unionized. This will be the case so long as there is only one union and the size of the union is increasing or in the steady-state, since unions always target the easiest to organize firms first.[7] In a given instant, there will be two segments of non-unionized firms, a "thin" segment of firms that have just been created with difficultly distributed according to the initial distribution and a "thick" segment of pre-existing firms with difficulties greater than p. Unions will optimally spend their organizing budget first to organize newly emerged firms in the thin segment with organizing difficulty below p. Once the union has organized those firms, it will spend what remains of its budget on the remaining previously existing firms in the thick segment with marginal difficulty of organizing p.[8]

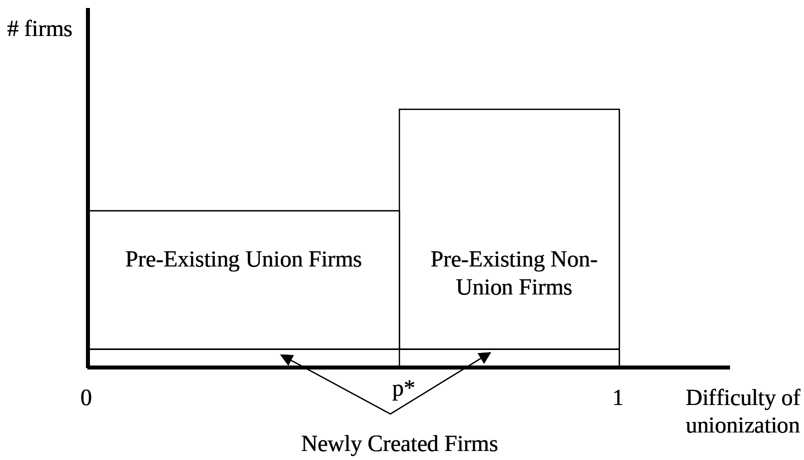
[6] One key difference between a model of unions and standard epidemiological models of disease, such as the SI model, is that as disease prevalence increases in an epidemiological model, the efficiency with which infected hosts pass on the disease declines. This occurs because standard epidemiological models assume random matching between hosts in the population. When the disease becomes very prevalent, many of these matches occur between two infected hosts, so those matches do not contribute to the spread of the disease. To the extent that unions spread through random word of mouth connections, similar dynamics would arise.

In this model, however, we assume that unions will not waste resources attempting to unionize firms which have already been unionized. Rather, they will explicitly target firms that are not yet unionized. An interior solution for union coverage arises because of heterogeneity among firms in the difficulty of union organizing.

[7] During transitions that involve the decline of a union–for example, in response to some kind of shock that reduces the union's effective organizing budget–there will actually be a range of costs where there will be both unionized and non-unionized firms. This will be discussed in more detail in the Appendix. In the steady state, however, there will be some p below which all firms are organized and above which no firms are organized.

[8] Strictly speaking, this suggests that in the steady state, the percentage of unionized firms will be higher among newly-created firms than among older firms. In practice, factors outside the model will obscure this relationship. For example, if firms differ in intrinsic profitability, more profitable firms will be more attractive to unions and longer-lived. To take another example, if firms take time to grow and initially face a high death rate, unions may not organize early in the firm's life.

Figure 1: Steady-State Unionization



A graphical depiction of the steady-state is given in Figure 1. Note that the density of unionized firms is lower than the density of non-unionized firms in steady-state, because although the cost distribution for newborn firms is uniform, unionized firms have a higher death rate, and therefore do not live as long as non-union firms.

Normalize the number of firms, F, to 1, so that U becomes the fraction of firms that are unionized. At an instant of time dt, $[\delta(\alpha)U + \delta(0)(1-U)]\,dt$ firms will have just exited due to a negative productivity shock. As those firms die, new firms will be born with difficulties of being unionized distributed according to the initial distribution. For a union to organize all newborn firms with difficulty level below p, the union will have to spend

$$[\delta(\alpha)U + \delta(0)(1-U)]\,dt \int_0^p c\; dG(c)\,, \tag{3}$$

which, since $G(c)$ is Uniform[0,1], is just

$$[\delta(\alpha)U + \delta(0)(1-U)]\,dt \frac{p^2}{2}. \tag{4}$$

In order for p, the threshold below which all firms are organized, to remain constant, the union's effective organizing budget must exactly correspond to the total cost of organizing all newly created firms with cost less than or equal to p, i.e.:

$$A(\alpha)BU = [\delta(\alpha) U + \delta(0)(1-U)] \frac{p^2}{2}. \tag{5}$$

This condition, that p must not change, is one of the two conditions that must be satisfied in the steady-state. If the union had a surplus, i.e. if $A(\alpha)BU > [\delta(\alpha) U + \delta(0)(1-U)] \frac{p^2}{2}$, then it would spend that surplus organizing non-union firms in the "thick" segment with difficulty greater than p, and p would increase. Conversely, if the union's budget was not sufficient to organize all of the newly born firms with difficulty below p, then p would decrease. (See Appendix A.2 for more details.)

The other condition that must be satisfied in the steady-state is that the number of unionized firms, U, must also not change. This means that the number of newly born firms the union organizes must exactly equal the number of firms the union loses to attrition. This yields the condition

$$[\delta(\alpha) U + \delta(0)(1-U)] p = \delta(\alpha) U. \tag{6}$$

These two conditions, that the difficulty distribution of unionized and non-unionized firms does not change and that the number of unionized firms does not change, lead us to the following characterization of the steady-state:

Proposition 1 *With a single union, there can be two steady-states, the trivial steady-state with no unionization ($U = p = 0$) and the steady-state with*

$$U^* = \begin{cases} \frac{2\delta(0)A(\alpha)B}{\delta(\alpha)^2 - 2A(\alpha)B[\delta(\alpha)-\delta(0)]} & \text{if } 2A(\alpha)B \leq \delta(\alpha) \\ 1 & \text{otherwise} \end{cases}; \tag{7}$$

$$p^* = \begin{cases} \frac{2A(\alpha)B}{\delta(\alpha)} & \text{if } 2A(\alpha)B \leq \delta(\alpha) \\ 1 & \text{otherwise} \end{cases}. \tag{8}$$

Proof. Equations (7) and (8) can be obtained by combining equation (5) and equation (6).

The derivation for the condition that guarantees an interior solution,

$$2A(\alpha)B \leq \delta(\alpha), \tag{9}$$

can be seen by setting the algebraic expressions for U^* and p^* equal to 1, the maximum value they can take, given that the maximum proportion of firms that can be unionized is 1 and that the difficulties of unionization are distributed on the interval [0,1]. ■

Note that when $2A(\alpha)B > \delta(\alpha)$, the union's organizing budget is substantial enough to overcome the attrition of member unions, so the model would be at a corner solution with steady state unionization levels of either 0 or 1. For the remainder of the paper we will assume that condition (9) holds unless otherwise stated, so that we are in the more interesting interior case with only partial unionization in the non-trivial steady-state.

As shown in Appendix A.3, the trivial steady-state with $U^* = 0$ is unstable, and the non-trivial steady state with partial unionization is stable. The intuition behind these results is that the total resources available for union organizing rise linearly with the number of unionized firms, while the cost of replacing firms lost to attrition rises faster than linearly given that the easiest firms to unionize are unionized first. Given our assumption of a uniform distribution of difficulty of unionization, the cost of replacing firms lost to attrition is quadratic in the level of unionization. The cost of replacing unionized firms lost to attrition is less than the resources available for unionization at all unionization levels between 0 and the non-trivial steady-state, and greater than the available resources curve at higher levels of unionization. With a non-uniform cost distribution, there could be multiple stable non-trivial equilibria, but we focus on a simple case here.

Since the distribution of unionization difficulties is uniform on [0, 1], p^*, the difficulty level below which all newborn firms are unionized, is also the percentage of newborn firms that are unionized. Combining condition (9) with the fact that $\delta(\alpha) \geq \delta(0)$ implies that $U^* \leq p^*$ (to see this, compare the equations for U^* (Equation (7)) and p^* (Equation (8)). Intuitively, U^*, the steady-state proportion of firms that are unionized, is less than p^*, the proportion of newborn firms unionized, because unionized firms die at a faster rate than non-union firms.[9] As can be seen in

[9] Of course, in the real world, factors outside the model may obscure this relationship. In particular, firms may

equation (7), the greater the difference between $\delta(\alpha)$ and $\delta(0)$, the more of the union's organizing budget it has to spend to make up for firms lost to attrition, and therefore the lower the steady-state value of U^*. On the other hand, when $\delta(\alpha) = \delta(0)$, U^* and p^* are identical.

Exogenous increases in the death rate of firms that increase $\delta(\alpha)$ and $\delta(0)$ by the same proportion will reduce steady-state unionization. The intuition is that with higher attrition rates, at every level of membership the union must devote a greater share of its resources to replacing firms lost to attrition and less to expanding the size of the union. This result is stated more formally in the following proposition and tested empirically in Section 5.

Proposition 2 *Increasing the death rate of all union and non-union firms by the same proportion reduces the steady-state level of unionization U^*.*

Proof. Suppose that the ratio $\frac{\delta(\alpha)}{\delta(0)}$ is fixed at $\lambda(\alpha)$. Then equation 7 can be rewritten as

$$U^* = \frac{2A(\alpha)B}{\lambda(\alpha)^2\delta(0) - 2A(\alpha)B[\lambda(\alpha) - 1]}. \tag{10}$$

Taking the derivative with respect to $\delta(0)$ yields

$$\frac{dU^*}{d\delta(0)} = -\frac{U^*\lambda(\alpha)^2}{\lambda(\alpha)^2\delta(0) - 2A(\alpha)B[\lambda(\alpha) - 1]}. \tag{11}$$

Condition (9) guarantees that $2A(\alpha)B \leq \lambda(\alpha)\delta(0)$, which in turn guarantees that $\frac{dU^*}{d\delta(0)}$ will be less than zero. ■

3 Rent Extraction Under Optimizing and Evolutionary Models

This section contrasts economic and evolutionary analyses of the determination of the level of rent extraction, α. Under a standard economic approach, unions choose α to maximize the present discounted value of rents to union members, taking into account the dependence of firm investment on α. Under the evolutionary approach, unions are endowed with different values of α, and only

differ in intrinsic profitability, and more profitable firms are more likely to attract attention from unions and less likely to exit.

those unions with evolutionarily stable values of α survive. In many circumstances, the economic and evolutionary approaches yield the same steady-state predictions, albeit with different dynamics (as in Nelson and Winter, 1982). In this model, however, the evolutionarily stable value of α will be less than the value of α that maximizes the present discounted value of rents to current union members.

The remainder of this section is organized as follows. Subsection 3.1 derives the conditions for the optimal level of α for the workers. Subsection 3.2 then shows that the evolutionarily stable level of rent extraction is less than this welfare maximizing level.

3.1 Welfare-Maximizing Level of Rent Extraction

We first consider a fairly conventional model in which unions choose α to maximize the present discounted value of rents accruing to current union members.

We assume that unions cannot commit to a path of rent extraction over time. Otherwise, the optimal contract would involve a one-time payment from the firm in exchange for an agreement to never again extract any rents. This would avoid distorting the firms's investments in staying alive. In fact, it is difficult to contract on rent extraction, since firms may not be able to specify in advance the exact tasks needed later and unions may have difficulty committing never to extract rents.

Given this, the union chooses how much it will extract each year. Since firms' pre-union profits are constant, there is no difference between extracting a lump sum each year and a share of profits each year. We will consider the case in which unions have all the bargaining power in negotiations with firms, in the sense that they can present firms with take-it-or-leave-it offers. This assumption may be reasonable if a single union bargains with many firms and has incentives to acquire a reputation for toughness. Although unions cannot commit to a time-path of future rent extraction, bargaining is statically efficient, so that all firms employ the efficient number of workers.

Suppose workers face a Poisson hazard rate ϕ of separation from the firm, for example through death or retirement. The present discounted value of rents accruing to current union members is

$$\frac{\alpha\pi}{r+\phi+\delta(\alpha)}. \tag{12}$$

Since $\delta(\alpha)$ increases with α, the optimal level of rent extraction for the worker involves a trade-off between the flow of rents and the hazard rate that the firm will chose, which would cause workers to cease to obtain any rents.[10]

The first order condition for the level of α that maximizes the present discounted value of rents for workers, denoted α_W, is

$$r+\phi+\delta(\alpha_W)-\alpha_W\delta'(\alpha_W)=0. \tag{13}$$

For the remainder of the paper, we assume that the parameter values are such that we have an interior solution for α_W.

3.2 Evolutionarily Stable Rent Extraction

An alternative approach to understanding how α is determined is to assume that α, the level of rents a union extracts, is fixed for a given union, but that there are many unions with differing levels of α. One can then ask which union will survive in evolutionary competition.

We assume that the function $A(\alpha)$, which indexes how attractive a union is to potential new members, is continuously increasing in the present discounted value of rents obtained by workers (i.e., equation (12)). The assumption that $A(\alpha)$ is increasing in the present discounted value of rents extracted by workers implies that a union that maximizes the welfare of its members, i.e. a union that extracts α_W, has the easiest time organizing unorganized firms.[11]

[10] Note that this expression assumes implicitly that workers receive no union rents if they leave the firm. This will be true if the labor supply, N, is large enough, so that the probability the worker obtains a second job in the potentially unionizable sector, and therefore has a chance of getting a unionized job, approaches 0.

[11] In fact, while we assume that $A(\alpha)$ is maximized at α_W, it is plausible that it is maximized at some value less than α_W. Firms can employ a wide variety of anti-unionization tactics, including requiring workers to attend anti-union meetings on company time, challenging the proposed definition of the bargaining unit, and illegally firing union activists, and the more they expect unions to extract, the more vociferously they will oppose unions. Given the response of firms' unobservable investment to α, as α approaches α_W, increases in α hurt firms much more than they help workers. Firms' opposition to unionization might therefore increase more rapidly with α than workers' support for unionization. Firms may even ease the entry of more moderate unions to forestall more radical alternatives. Such effects, however, would only make showing that $\alpha_S < \alpha_W$ easier, so we ignore any effects of this sort in the model.

The assumption that A is continuous in α is important for the result that $\alpha_S < \alpha_W$. If there were simple Bertrand competition among unions for potential members at unorganized firms, in which workers joined whichever union delivered greater discounted rents, then the evolutionarily stable level of rent extraction would equal the optimal amount of rent extraction for current workers. However, if workers decide which union to join based not only on the present discounted value of rent extraction but also on other idiosyncratic factors, such as the match between the personality of union organizers and the workers at the firm, then workers may join a union other than the one that maximizes the present discounted value of rents. Union recruitment will therefore increase continuously rather than discretely in the present discounted value of rents delivered to members.

If there are multiple unions, each would like to spend its organizing budget trying to organize the easiest firms. Rather than assume that unions waste resources on battles to organize the same unorganized firms, we will assume that they divide them so that at every level of difficulty, c, unions organize firms in proportion to their effective organizing budgets. Since the effective organizing budget is the actual organizing budget (BU) multiplied by how attractive the union is to workers (indexed by the function $A(\alpha)$), unions that are more attractive to workers can organize disproportionately more firms.[12] For example, suppose that there are two unions, a moderate union with M member firms and extraction rate α_M and a radical union with R member firms and extraction rate α_R. The moderate union targets $\frac{A(\alpha_M)BM}{A(\alpha_M)BM+A(\alpha_R)BR}$ of the non-unionized firms with difficulty less than p and the radical union targets the remainder.

We can now identify the evolutionarily stable level of rent extraction and show that it will be smaller than the welfare-maximizing level of rent extraction. First, we specify how the definition of evolutionary stability applies in our context.

Definition 1 *A union that extracts a rent level α is* evolutionarily stable *if and only if, starting from the steady state containing only the α union, there exists an $\gamma > 0$ such that if any other union with size $\varepsilon < \gamma$ invades, the invading union will disappear.*

[12] We thus allow for unions that extract more for their members to be more successful in attracting members, but rather than having a completely general function for union recruiting as a function of the union's level of rent extraction and that of each of its competitors, we consider the case in which each union's recruiting is proportional to its attractiveness to workers and its organizing budget.

Proposition 3 *The union that extracts the level of rent α that maximizes the ratio $\frac{2A(\alpha)B}{\delta(\alpha)}$ will be evolutionarily stable.*

Proof. Denote by α_S the level of α that maximizes $\frac{2A(\alpha)B}{\delta(\alpha)}$. Let S represent the number of unionized firms in the union that extracts α_S. Consider a steady-state containing only the α_S union, and introduce into this steady-state a small union of size $\varepsilon > 0$ that extracts $\alpha_\varepsilon \neq \alpha_S$. In order to show that α_S is evolutionarily stable, we need to show that for each α_ε, there exists a minimum size γ such that if the size of the invading union ε is less than γ, then the invader will have negative growth and die off. To see that this will be the case, consider how the ε union spends its effective organizing budget of $A(\alpha_\varepsilon) B\varepsilon$. With such a budget, it can afford to organize the newborn firms up to some level p_ε, determined by setting the effective organizing budget equal to the number of newborn firms times the proportion organized by the invading union times the average cost of unionization for firms with cost less than p_ε:

$$A(\alpha_\varepsilon) B\varepsilon = [\delta(\alpha_S) S + \delta(\alpha_\varepsilon)\varepsilon + \delta(0)(1 - S - \varepsilon)] \frac{A(\alpha_\varepsilon) B\varepsilon}{A(\alpha_S) BS + A(\alpha_\varepsilon) B\varepsilon} \int_0^{p_\varepsilon} c\, dc, \qquad (14)$$

which yields

$$p_\varepsilon = \sqrt{\frac{2[A(\alpha_S) BS + A(\alpha_\varepsilon) B\varepsilon]}{\delta(\alpha_S) S + \delta(\alpha_\varepsilon)\varepsilon + \delta(0)(1 - S - \varepsilon)}}. \qquad (15)$$

Recall from the single-union case (equations (5) and (8)) that in the steady state,

$$p_S^* = \frac{2A(\alpha_S) B}{\delta(\alpha_S)} = \sqrt{\frac{2A(\alpha_S) BS}{\delta(\alpha_S) S + \delta(0)(1 - S)}}. \qquad (16)$$

Note that when ε is close to 0, p_ε is approximately equal to p_S^*. Since in the steady state before the invasion all firms with difficulty level less than p_S^* are unionized, when ε is close to 0 the invading union will exhaust its budget organizing firms up to p_ε. The growth rate of the invading union will be

$$\dot{\varepsilon} = [\delta(\alpha_S) S + \delta(\alpha_\varepsilon)\varepsilon + \delta(0)(1 - S - \varepsilon)] \frac{A(\alpha_\varepsilon) B\varepsilon}{A(\alpha_S) BS + A(\alpha_\varepsilon) B\varepsilon} p_\varepsilon - \delta(\alpha_\varepsilon)\varepsilon. \qquad (17)$$

Rearranging terms, we find that the growth rate of the invading union $\dot{\varepsilon}$ will be less than 0 if

$$\frac{2A(\alpha_\varepsilon)B}{\delta(\alpha_\varepsilon)} < \sqrt{\frac{2\left[A(\alpha_S)BS + A(\alpha_\varepsilon)B\varepsilon\right]}{\delta(\alpha_S)S + \delta(\alpha_\varepsilon)\varepsilon + \delta(0)(1 - S - \varepsilon)}}. \tag{18}$$

Since the RHS equals p_ε and p_ε can be made arbitrarily close to p_S^* by setting ε small enough, we can re-write this inequality as

$$\frac{2A(\alpha_\varepsilon)B}{\delta(\alpha_\varepsilon)} < \frac{2A(\alpha_S)B}{\delta(\alpha_S)}. \tag{19}$$

Since the ratio $\frac{2A(\alpha)B}{\delta(\alpha)}$ is precisely what α_S maximizes, we know that this inequality will hold and that the α_S union will be evolutionarily stable. ■

The key idea of the proof is that $\frac{2A(\alpha)B}{\delta(\alpha)}$ is the steady-state level of p^*, the proportion of newborn firms that are unionized in steady-state. This determines the average cost level the union can sustain in steady-state. A union that can bear a higher average cost level than the incumbent will be able to unionize disproportionately more firms, and will be able to invade; a union unable to bear as much will experience negative growth and disappear. Therefore, no union can successfully invade a steady-state containing the union with the highest possible average cost level. The union with the maximum value of $\frac{2A(\alpha)B}{\delta(\alpha)}$ is therefore evolutionarily stable.

It is worth noting that while the level of α that maximizes $\frac{2A(\alpha)B}{\delta(\alpha)}$ will be unique under most normal parameterizations of $A(\alpha)$ and $\delta(\alpha)$, this need not hold in general. It is possible to construct functions $A(\alpha)$ and $\delta(\alpha)$ satisfying all of the conditions above such that $\frac{2A(\alpha)B}{\delta(\alpha)}$ has multiple global maxima. In this case, there will be several possible levels of rent extraction α_S that, together or independently, would be evolutionarily stable. However, finding examples of functions $A(\alpha)$ and $\delta(\alpha)$ satisfying all of the conditions above and where $\frac{2A(\alpha)B}{\delta(\alpha)}$ has multiple global maxima requires careful construction, so it seems likely that this will not occur empirically.

Proposition 3 guarantees that, starting from a steady state occupied by only the α_S union, no other union can invade. We now show that facing a steady state containing any other union or combination of unions, the α_S union can successfully invade. Furthermore, if the system then converges to a steady state, that steady state will contain only the α_S union. To show this, it will be useful to first state the following lemma.

Lemma 1 *If multiple unions co-exist in the steady-state, then they must have the same ratio of effective organizing budget to firm death rate, i.e.* $\frac{2A(\alpha_M)B}{\delta(\alpha_M)} = \frac{2A(\alpha_R)B}{\delta(\alpha_R)}$. *Furthermore, this ratio will be equal to the organizing cost of the most difficult to organize firm that is unionized in the steady-state, i.e.* $p_M^* = \frac{2A(\alpha_M)B}{\delta(\alpha_M)}$.

The proof is given in Appendix A.1. The intuition behind the Lemma is that for two unions to exist in the steady-state, one must be a more moderate union that is less attractive to workers but loses fewer of its member firms due to attrition, while the other must be a more militant union that is better able to unionize new firms but also loses more of its member firms to attrition. Lemma 1 specifies how precisely to balance this trade-off.[13]

With this lemma characterizing the steady-state in mind, we can show that an evolutionarily stable union will be able to invade any other union.

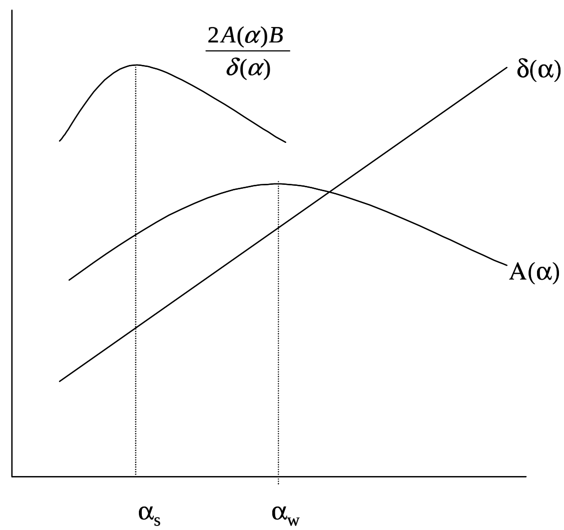
Proposition 4 *The* α_S *union can successfully invade any steady-state other than the one containing another* α_S *union.*

The proof is essentially similar to the proof of Proposition 3, and is given in Appendix A.1.

We have so far shown that, starting from a steady-state containing the α_S union, no union can invade, and starting from a steady-state with any other union, the α_S union can invade and grow. We have not ruled out a limit cycle, but we do know that if there is a steady-state, it must be the steady-state containing only the evolutionarily stable union. To see this, suppose that there are two unions, the stable union S and an incumbent union I. By Lemma 1, the eventual steady state cannot contain both the S union and the I union, since they have different ratios $\frac{2A(\alpha)B}{\delta(\alpha)}$. We have already shown that as the world approaches the steady-state with the I union, whatever tiny amount ε of the S union that remains will grow, so the ε union can not be eliminated entirely. Therefore, we have shown that the α_S union cannot be displaced by any other union and that, assuming that there is no cycling, the α_S union can invade and displace any other union.

[13] Note that when the function $\frac{2A(\alpha)B}{\delta(\alpha)}$ is strictly concave, which it will be for many (but not all) concave functions $A(\alpha)$ and increasing functions $\delta(\alpha)$, there can be at most two unions in equilibrium. When $\frac{2A(\alpha)B}{\delta(\alpha)}$ is not strictly concave, on the other hand, there can be three or more unions in equilibrium. Even in this case, however, the same argument in Lemma 1 goes through.

Figure 2: Evolutionarily stable and welfare-maximizing levels of rent extraction.



Now that we know which union will be evolutionarily stable, we can show our key result: that it is less radical than the welfare-maximizing union.

Proposition 5 *The evolutionarily stable level of rent extraction, α_S, is smaller than the level of rent extraction that maximizes the present discounted value of wages of current members, α_W.*

Proof. As shown above, the evolutionarily stable level of rent extraction, α_S, maximizes the ratio $\frac{2A(\alpha)B}{\delta(\alpha)}$. Since α_W, the level of rent extraction that maximizes the present discounted value of wages of current union members, maximizes $A(\alpha)$, and since δ monotonically increases in α, $\frac{2A(\alpha)B}{\delta(\alpha)}$ is decreasing in α at α_W and at all greater values of α. Since α_S maximizes $\frac{2A(\alpha)B}{\delta(\alpha)}$, it must be less than α_W. ■

Figure 2 presents the proof graphically, showing $A(\alpha)$, $\delta(\alpha)$, and $\frac{2A(\alpha)B}{\delta(\alpha)}$ as functions of α. δ increases monotonically with α, and $A(\alpha)$ increases with α up to α_W, the level of output that maximizes the welfare of current workers, and then declines. This implies that α_S, the evolutionarily stable level of rent extraction, is less than α_W. If one starts at the level of rent extraction that is optimal for members, a small reduction in α causes a second-order reduction in attractiveness of the union to potential members, and thus a second-order reduction in the spread rate of the union. However, it causes a first-order decrease in the exit rate of unionized firms. Therefore, the evolutionarily stable level of α must be less than the welfare-maximizing level of α. This result holds as long as the spread rate of unions is continuous in the present discounted value of wages extracted.

Note that the relative shapes of the $A(\alpha)$ and $\delta(\alpha)$ functions determines how far α_S will be from α_W. If $A(\alpha)$ declines gradually as one moves away from α_W, then α_S is likely to be considerably less than α_W. On the other hand, if $A(\alpha)$ declines steeply as one moves away from the welfare maximizing level of output, then α_S will be very close to α_W. Moreover, if $\delta(\alpha)$ is steep, so that firm survival is sensitive to rent extraction, then α_S will be far below α_W, whereas if $\delta(\alpha)$ is fairly flat, then α_S will be close to α_W.

The steady-state number of unionized firms in society is higher if unions extract α_S than if they extract α_W. Furthermore, the level of rent extraction that maximizes the number of unionized firms will be less than or equal to the evolutionarily stable level α_S, and therefore, by Proposition 5, less than α_W. The intuition for this result is that, under any union more radical than the α_S union, a smaller percentage of newly created firms are unionized (since α_S maximizes p^*) and the death rate of those firms is higher (since $\delta(\alpha)$ increases monotonically with α). The following proposition shows these results formally.

Proposition 6 *For any level of rent extraction α greater than the evolutionarily stable level of rent extraction α_S, the steady-state level of unionization, $U^*(\alpha)$, will be lower than the steady-state level of unionization under the evolutionarily stable union, $U^*(\alpha_S)$. This implies that the level of α that maximizes the steady-state level of unionization will be less than or equal to the level that maximizes α_S.*

The proof is given in the Appendix.[14]

Corollary 6.1 *The steady-state level of unionization under the evolutionarily stable level of rent extraction, $U^*(\alpha_S)$, will be greater than that under the welfare-maximizing level of rent-extraction, $U^*(\alpha_W)$.*

[14] Note that technically, if the function $\frac{2A(\alpha)B}{\delta(\alpha)}$ has multiple global maxima, so that the evolutionarily stable level $\{\alpha_S\}$ is not unique, then this result holds for the highest α_S belonging to that set.

Proof. This follows directly from Proposition 6 and from the fact that $\alpha_S < \alpha_W$, which was shown in Proposition 5. ■

It is worth noting, however, that it is ambiguous whether the total flow of rent extracted by the union in steady-state, $U^* \alpha \pi$, would be higher or lower with the evolutionarily stable union than with the welfare-maximizing union.[15] The reason is that decreasing α from α_W to α_S increases the number of unionized firms, but decreases the amount extracted from each firm. It is theoretically ambiguous which of these two effects dominates.

4 Implications for Union Institutions

Much as the characteristics of an organism are determined by its DNA, union behavior is influenced by its constitutional provisions. The model predicts that unions pursuing the wage policies that would be preferred by members could be displaced in evolutionary competition with unions that moderate these policies. But in the absence of incumbency advantages, union officials pursuing such moderate policies would lose office. The model thus suggests that unions with constitutional incumbency advantages may potentially be able to displace other unions in evolutionary competition. Section 4.1 provides examples of how union institutions are structured to give incumbency advantages to union leaders. Section 4.2 then provides evidence that when these incumbency advantages are weaker, union wage demands tend to become less moderate.

4.1 Incumbency Advantages

Incumbents have a substantial advantage over their potential challengers in union elections. Table 1 shows the turnover of union presidents for the ten largest U.S. unions since each union's founding. We focus on the chance an incumbent was defeated each year as a measure of incumbency advantages, since this captures both the advantages incumbents have through infrequent elections

[15] To see this, consider the following simple example. Suppose for simplicity that the discount rate r is equal to 0, and normalize the flow of profits π to 1. Suppose that the function $A(\alpha) = 1 - (\alpha - b)^2$, where $b \in (0, 1)$. Suppose that the function $\delta(I) = \frac{1}{I}$. This implies that $\delta(\alpha) = \frac{2}{(1-\alpha)}$, that $\alpha_W = b$, and that $\alpha_S = \frac{1+2b-\sqrt{3-(1-b)^2}}{3}$. In this case, when b is less than approximately .62, $\alpha_S U^*(\alpha_S) > \alpha_W U^*(\alpha_W)$, whereas when b is larger, the inequality is reversed.

Table 1: Turnover of union presidents for 10 largest American unions.

Union		Year Founded	Total Number of Presidents	Average Tenure of Presidents	Number of Defeated Incumbents	Chance Incumbent Defeated Per Year
1. National Education Association (NEA)		1934	12	5.5	3	4.5%
2. Teamsters (IBT)		1903	6	16	2*	2.1%
3. Food & Commercial Workers (UFCW)		1912	14	6	0	0.0%
4. State, County Employees (AFSCME)		1932	3	22	0	0.0%
5. Teachers (AFT)		1916	15	6.5	2	2.4%
6. Auto Workers (UAW)		1947	8	6.5	1	1.9%
7. Electrical Workers (IBEW)		1890	16	7	2	1.8%
8. Communication Workers (CWA)		1938	3	21	1	1.6%
9. Machinists (IAM)		1888	13	9	N/A	N/A
10. Steelworkers (USW)		1894	6	18	0	0.0%
Average:	All Unions		9.6	11.8	1.2	1.3%
	Private Sector Unions**		8.6	13.2	0.9	1.0%
	Public Sector Unions**		10.0	11.3	1.7	2.3%
Comparison:						
Presidents of the United States (1900-2000)			18	5.6	5***	5.0%

Source: National union offices.

* Both of the defeated Teamsters presidents were defeated after the Federal government takeover of the union and the imposition of direct elections for the union president.
** The NEA, AFT, and AFSCME are classified as public sector unions; the remainder are classified as private sector unions.
*** General election defeats.

and the electoral advantages gained once an election is held. Over the history of the nine unions for which data is available, an incumbent union president had only a 1.3% chance of being defeated in an election each year. This figure would be even lower if one excludes two defeats in the Teamsters union, which both came after the Federal government took over the union and imposed changes in election procedures that decreased incumbency advantages. To put these numbers in perspective, during roughly the same period, a President of the United States had a 5% annualized chance of being defeated in an election.

This data seems generally consistent with the model. In general, the relatively low chance of an incumbent being defeated is consistent with the high levels of incumbency advantages necessary to sustain more moderate wage policies.

Furthermore, note that the two unions with the greatest chance of an incumbent being defeated each year were the National Education Association and the American Federation of Teachers, both public sector unions. More broadly, incumbents in public sector unions had a 2.3% chance of defeat, while those in the private sector had a 1% chance of defeat (0.8% excluding the two Teamster defeats.) The model suggests that incumbency advantages are less important in public sector firms. To see this, note that there is essentially zero chance that a public employer will be forced out of business. The function $\delta(\alpha)$ is therefore much flatter than for private sector firms.[16] As a result, α_S will be very close to α_W, and weak incumbency advantages will not lead to much selective pressure against a union. The fact that the chance of a union president being defeated in public sector unions is almost twice that of private sector unions is consistent with this prediction, though there are other possible explanations as well.

One reason why incumbents are so often reelected is that most existing unions have constitutional features creating substantial incumbency advantages for leaders.[17] For example, most unions have indirect leadership elections, in which the president of the union is elected by delegates to a national convention, rather than by the membership at large. At these conventions, the delegates, often local union leaders, face strong pressure to support incumbents in national office if they think that the incumbents will win, because local union leaders need several types of services from national unions. For example, the union leadership often controls access to national strike funds and has the power to put local branches in trusteeship. [Geoghegan, 1992; Benson, 1986]. Furthermore, since incumbency advantages are much weaker in some union locals than at the national level, local leaders face the threat of not being re-elected and having to return to the shop floor. Local leaders' insurance against this threat is the possibility of obtaining a job with the national union staff, which will be much more likely to occur if they have reliably supported the national leadership. All of these factors encourage the delegates to the national conventions to support the incumbents.

Union incumbents have other direct advantages over challengers as well. Union staff are often

[16] The function $\delta(\alpha)$ is probably not completely flat, as militant actions on the part of unions can provoke a government to de-unionize. One classic example of this is President Reagan's confrontation with the air traffic controllers. Such situations are, however, relatively rare.

[17] Lipset, Trow, and Coleman [1956] examine the special case of the International Typographical Union, which had a functioning two-party system within the union, and conclude that outside of this special case, there was little prospect for true worker control of unions.

not restricted from donating money to support campaigns of current leadership, and laws restricting union staff from campaigning on union time are extremely weak. To take another example, union officers are not often required to give membership lists or even lists of local chapters to opposition candidates. Since unions often represent diverse sets of workers (for example, the United Auto Workers represents graduate students at NYU), this makes it difficult for challengers to campaign against incumbent leaders. On the other hand, incumbents can use official union communications, such as union newsletters, to promote their own candidacies. Even if there is a viable challenger, local union officers, rather than neutral third parties, are typically in charge of vote counting in union elections [Geoghegan, 1992], so there are few safeguards against fraud. In fact, there is anecdotal evidence of a significant amount of outright vote-stealing in union elections. Moreover, prior to mandated periodic elections under federal law, unions could go for decades without even holding elections. For example, the Laborers' union had no conventions between 1920 and 1941 [Benson, 1986].

While insulating union leaders from their membership through incumbency advantages allows leaders to use their power to moderate workers' wage demands, it does not guarantee that they will do so. There are, however, several reasons why union leaders may prefer lower wage demands. First, just as firm managers are often assumed to be empire builders, with a preference for increasing firm size, union leaders may prefer to be in charge of larger unions, as leaders of larger unions have more prestige and political power. As was shown in Proposition 6, increasing the steady-state size of the union requires extracting less than the welfare-maximizing level of rent. Therefore, union leaders that care about prestige will extract less rent than would be preferred by their members.

Moreover, in return for favors from firms, union leaders may collude with managers to moderate the union's wage demands. In fact, Ross [1950] argues that unions are often prepared to sacrifice worker-oriented provisions, such as wages, for union-oriented provisions, such as union security, automatic checkoff of union dues, the right of the union to participate in all grievance negotiations, and preferential seniority for union officials. Union leaders can thereby improve their own position by moderating workers' wage demands.[18] If the surplus from collusions is split between union

[18] One possible objection is that, incumbency advantages notwithstanding, union leaders are constrained by the requirement that most contracts must be approved by a majority of the rank-and-file. However, because the union

leaders and firms, then this collusion will allow firms to live longer. Of course, if incumbency advantages are too strong, union leaders wind up taking enough that workers will not want to join the union in the first place.

Finally, note that in an evolutionary model, there need be no presumption that all union constitutions that create incumbency advantages also create incentives for moderation. If some union constitutions create incumbency advantages but have provisions that encourage leaders to be more radical than members would prefer, these unions will die out. Meanwhile, if other union constitutions create incumbency advantages and also encourage leaders to moderate members' wage demands, these unions will grow.

4.2 Comparative Statistics of Wage Demands and Incumbency Advantages

If union leaders are protected by incumbency advantages and use that protection to moderate workers' wage demands, as the model predicts, then incumbents should favor more moderate wage demands than outsiders, and in those cases where incumbency advantages are weaker, unions should have stronger wage demands.

There is evidence for both predictions. Evidence that union leaders typically favor more moderate policies than would be preferred by members comes from the asymmetry of challenges to established leaders. Union dissidents typically accuse union leaders of being too moderate in their negotiations with the firms, not of threatening members' jobs by being too radical. If union leaders sought to represent the typical worker, one would expect challenges to come as often from either direction.

There is also evidence that stronger incumbency advantages are associated with more moderate wage policy. While incumbency advantages are strong at the national or international level, union locals vary in the degree of control of incumbency advantages, and in some union locals, there is regular turnover of leadership. We would therefore expect that the weaker the incumbency

leadership controls agenda setting, proposing contracts which rank-and-file can only approve or reject, union leaders have considerably more bargaining power than the rank-and-file. A simple model of such a bargaining game suggests that the union leadership has full bargaining power and obtains their most preferred outcome. (See Appendix A.4 for a proof.) In practice, when contract proposals are rejected by the rank-and-file, union leadership often simply repackages the contract in new language rather than fundamentally altering the contract offer.

advantages in the local, the more militant that local will be. Kleiner and Pilarski [2001] find exactly such an effect in a comparison of two similarly-sized locals of the UAW. One local, comprised of many plants spread out over the Los Angeles area, was organized by the UAW with indirect elections, whereas the second local, comprised primarily of a single large plant, was organized with direct elections of union officials. Kleiner and Pilarski found that the geographically concentrated local with direct elections had a much more vigorous union democracy and much more aggressive wage demands.

If locals have weaker incumbency advantages than national unions, we should also expect that local unions should advocate stronger wage demands than national unions. In fact, this is generally the case, and there are a number of examples of local unions conducting strikes against the wishes of the national union. For example, the P9 Hormel strike was conducted by the local union without the support of the national union, as was the recent Caterpillar strike.[19]

The model also suggests that if incumbency advantages decline exogenously, wages will rise and firms will be more likely to fail. It is instructive to examine a case study of two unions that for plausibly exogenous reasons were subject to shocks that reduced incumbency advantages. In the late 1930's, John L. Lewis, the president of the United Mine Workers (UMW) and founder of the CIO, feuded with Roosevelt, going so far as to endorse Wendell Willkie, Roosevelt's Republican opponent. As part of an effort to enhance his national political stature, Lewis, who faced no serious opposition within the UMW, instituted direct leadership elections. The Steelworkers, which were created by the UMW, adopted a similar constitutional provision.

By the 1970's, leadership of the UMW had passed to the corrupt Tony Boyle. Just after the 1969 leadership election, Boyle arranged for the murder of his opponent, "Jock" Yablonski, and of Yablonski's family. This over-reaching led to intense federal scrutiny of the 1972 UMW election and the victory of the challenger, Arnold Miller. Miller's victory was followed by much increased militancy on the part of the union, the decline of the Eastern coal industry, and a dramatic decline in union membership.

[19] It is not clear what other models would predict about the relative militancy of the national union and locals. On the one hand, the national has to provide resources to support the local union in strikes, for example through the strike fund. On the other hand, a national union might wish to demonstrate its willingness to strike against other employers by striking against one employer.

Following the election defeat of the incumbent UMW leadership, in 1977 a major challenge was also launched to the Steelworkers' leadership, which was similarly vulnerable due to its constitutional provision for direct leadership elections. Before the election, the heir apparent, Lloyd McBride, had promised to make a number of concessions to management in the hopes of saving jobs in the ailing steel industry. Ed Sadlowski, McBride's opponent, challenged McBride as being too close to management, and was explicit about his willingness to sacrifice union membership for higher wages. Sadlowski said that he did not mind if the Steelworkers' membership dropped from 400,000 to 100,000 or even 60,000, and that it should be a goal of labor to have the steel industry pay high wages that would allow its workers to finance education so that they or their children could obtain better jobs. It is hard to imagine typical incumbent union leadership adopting policies that would cut membership to a quarter of its initial level. Though Sadlowski lost the election, as a result of his challenge McBride was forced to drop his concessions to management and adopt much more aggressive wage demands. With several years, the steel industry had begun a precipitous decline, shedding 56 percent of its workforce in the period from 1979 to 1986, a decline from which it has yet to recover [Tornell, 1997]. Of course, the decline of the Eastern coal industry and the U.S. steel industry was probably the result of a number of other factors as well, but the model is at least consistent with the data.

Finally, unions in the U.S., where labor laws greatly favor incumbent unions, seem to be more moderate than many of their European counterparts, where the threat of entry by competing unions may prevent incumbent leaders from departing too far from the workers' preferred policies. (Of course, this does not apply to the same extent in countries with encompassing unions on the Scandinavian model, where unions may have other incentives to moderate wages.)

Relative to labor law in most of Europe, U.S. labor law enhances incumbency advantages for existing unions. In the US, once a particular union has won a union certification election, it is officially recognized as the sole collective bargaining partner representing the covered workers, and it can only be replaced if the majority of workers vote to decertify it and then certify another union. Decertification, however, is relatively rare. In some European countries, such as France, Italy, and the Netherlands, several different unions may compete for workers within the same firm on an ongoing basis. The threat of entry makes it more difficult for incumbents to depart from members' preferred policies.[20] Reducing rent extraction from the level that maximizes the present discounted value of rents for current union members may increase the lifespan of firms, but it will lead to the loss of workers within the firm to rival unions.

As a result, individual-level selection is likely to be a much more potent force in European countries with multiple unions inside a single firm than in a U.S.-style system in which a single union is certified to collectively bargain on behalf of a defined set of workers. Even in countries such as Britain, where a single union typically represents a given set of workers, the weakness of barriers to entry for competing unions relative to the U.S. means that the implicit threat of competition is likely to constrain unions to represent their members relatively well.

Evolutionary and standard maximizing models differ most sharply in their predictions of relative militancy of unions under the U.S. system of multiple craft unions representing different types of workers within the firm and European systems in which different unions can potentially compete for the same potential members. In the U.S. craft union system, for example, airline pilots, machinists, and flight attendants are all represented by separate unions, and hence under standard maximizing models, if there are many unions each union has no incentive to internalize the effect of its own rent extraction on the firm's investment. Standard maximizing models therefore imply that rent extraction should therefore be greater in this craft union environment than under a European environment in which multiple unions compete within a single firm but wage concessions to one union apply to all employees. In an evolutionary model, however, the ongoing competition for members among unions in the European system could lead to more rent extraction than under a system of U.S.-style craft unions. The model is consistent with the widespread view that European unions are more militant than their U.S. counterparts.[21]

[20] Ongoing within-firm competition for members among unions will produce lower long-run rent extraction than restricting competition to the initial choice of union. This is because if unions only compete at some initial stage, unions that initially extract the level of rents which maximizes the present discounted welfare of members, and then gradually lower rent extraction, will be able to attract members with a policy which approaches the evolutionarily stable policy in the long run. Note that this policy does not require a commitment technology for unions, because it does not involve promises to undertake time-inconsistent policies. Extra benefits to workers joining a union are provided in the short run, not the long run. For example, unions could make an up-front payment in the form of support for organizers and support for an initial strike if necessary. In contrast, unions must maintain a high level of rent extraction in the long run to retain members in the face of ongoing competition.

[21] More systematic evidence on relative rent extraction is hard to come by. Wage premia for union members as

A similar comparison can also be made within the U.S. Prior to the merger of the AFL and the CIO, unions affiliated with each of the two umbrella organizations often continually competed to organize a given set of workers. This higher level of competition seems to have coincided with more militant behavior on the part of unions, as the model would predict.

The analysis of how rent extraction differs depending on whether or not unions compete within firms is analogous to the analysis of the evolution of virulence in biology. The strength of selective pressures for organisms to become more benign or even symbiotic depends on the mode of transmission of the organism [Ewald, 1994] . For example, if several different HIV strains are competing within the human body, one that reproduces more rapidly within the human body may be more likely to kill its host, but will also be more likely to be transmitted to another host. Thus individual-level selection within the host favors rapid reproduction while group-level selection favors more benign forms of the disease that are less likely to kill the host. In contrast, mitochondria reproduce only through cell division, so selection among mitochondria favors those that help their cells survive. Similarly, the system of incumbency advantages built into U.S. labor law produces an advantage for unions that help their firms survive. The greater ongoing competition for members among several different unions, the more this effect is counterbalanced by the need to extract more rent to attract members.

5 Firm Turnover and Unionization

Proposition 2 implies that exogenous increases in firm turnover reduces steady-state unionization levels. Recall from equation (11) that

$$\frac{dU^*}{d\delta(0)} = -\frac{U^* \lambda(\alpha)^2}{\lambda(\alpha)^2 \delta(0) - 2A(\alpha) B [\lambda(\alpha) - 1]} < 0, \tag{20}$$

conventionally measured are higher in the U.S. However, the lower union coverage in the U.S. means that wage premia may not be a good measure of rent extraction. In the U.S., unions may only be present in industries and firms with large amounts of rents to extract, whereas in Europe, unions are widespread.

so increases in the turnover of firms reduce steady-state unionization rates. The model thus suggests that unions should be more prevalent in industries with low firm turnover.[22] [23] This section empirically tests this implication using data for U.S. manufacturing industries.

We test whether industries with a high turnover of firms have low unionization rates. We use data on exit rates across U.S. manufacturing industries taken from Dunne, Roberts and Samuelson's [1988] analysis of the Census of Manufacturing. Their data set covers all firms producing in each four-digit SIC manufacturing industry in the census years of 1963, 1967, 1972, 1977, and 1982. They measured exit rates between the census years for two groups of firms–one group with all firms present in the years of interest, and a second group that excludes the smallest firms from each industry. The excluded firms in the second group are chosen such that the firms excluded produced less than 1% of their respective industry output. EXIT and EXIT99 are respectively the average yearly exit rates over the period for all firms and for the firms accounting for 99% of output. We focus on the latter measure, but results are similar for the former.[24][25]

Data on union membership and coverage rates by industry are taken from Hirsch and Macpherson's [1993] analysis of the Current Population Survey (CPS). Their data set includes average union membership and coverage for 231 three-digit Census of Population industries over the 1983 to 1991 period. The number of observations from the CPS for each industry ranged from 5 to 21,950, with

[22] Susan Dynarski has pointed out that a similar process may operate on a micro-level as part of organizing a particular firm. Union supporters within a firm influence their friends to become union supporters. If there is a high turnover rate among workers, it is very hard to organize the firm.

[23] The equation for U^* only literally applies at an industry-by-industry level if unions organize only within their own industry. However, we expect that similar results would arise if unions were disproportionately likely to unionize within their own industry, as is the case empirically. Current union members will gain more by unionizing within their own industry, since this reduces pressure on their own wages. Moreover, unions probably have specialized knowledge of how to appeal to and negotiate on behalf of workers in particular industries. Moreover, even if unions randomly tried to unionize firms from any industry, industries with longer-lived firms would have a larger proportion of unionized firms.

[24] Note that if a firm that was producing in more than one four-digit industry stopped producing in one of these industries, that would be counted as an exit for that particular industry (even though the firm might continue producing in the other industries).

[25] The raw data gives 5-year exit rates, but we converted these to annual rates as follows. First, we calculated the average survival rate by multiplying the survival rates between pairs of census years (defined as one minus the exit rate between the two years) and raising it to the power $\frac{1}{19}$ (since the data covers a 19 year interval). Once we had constructed the average survival rate, the average exit rate was defined as the difference between one and the average survival rate. The formula for the exit rate is thus

$$ext = [1 - (1 - ext6367) * (1 - ext6772) * (1 - ext7277) * (1 - ext7782)]^{\frac{1}{19}}$$

where $extxxyy$ is the exit between census years xx and yy.

Table 2: Means and Standard Deviations of the Variables

Variable	Mean	Std. Dev.
UMEM (%)	26	12
UCOV (%)	28	12
EXIT	10.9	1.9
EXIT99	7.6	2.1
HHI	.07	.04
SIZE (employees/establishment)	98	143
K (US$1,000/worker)	42	50

a mean of 3,857 and a standard deviation of 4,319. Given the implied variation in the accuracy of the union membership estimates, we weight each industry by the square root of the number of observations from which the union membership average was constructed.

We control for several industry characteristics from the 1992 Census of Manufactures. Data was obtained for employment, number of establishments, capital, and the Herfindahl-Hirschmann index of concentration (HHI). The data covered four-digit SIC manufacturing industries. We defined average plant size (SIZE) as employment divided by the number of establishments, and capital intensity (K) as the ratio of total assets to total employment (measured in US$1,000 per worker).

To merge the data sets so the four-digit SIC codes would match the three-digit Census of Population industry codes, it was often necessary to combine several four-digit industries into one. When combining the four-digit SIC industries, the Herfindahl-Hirschmann index and the exit rates were weighted by the employment in that four-digit SIC industry. After combining the different data sets, we have data for union membership and coverage, exit rates, average plant size, a measure of capital intensity and the Herfindahl-Hirschmann index for 66 Census of Population industries. Table 2 shows summary statistics for this sample.

Across a variety of specifications, higher firm exit rates are associated with lower unionization rates. (See Table 3.) In our preferred specification, shown in column (5) of Table 3, a 1 percentage point increase in the exit rate is associated with a 3.4 percentage point decrease in the unionization rate. The results are fairly comparable, whether union membership (UMEM), or union coverage

Table 3: Union Membership and Coverage regressed on exit rates and other industry characteristics, with industries being weighted by the square root of the number of observations from the CPS.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Variable	UMEM	UMEM	UMEM	UMEM	UMEM	UMEM	UCOV
EXIT99	-3.54**	-3.06**	-3.94**	-3.53**	-3.44**		-3.38**
		(.66)	(.75)	(.78)	(.87)		(.88)
EXIT						-1.45*	
						(.856)	
HHI		60.65**			92.44**	145.68**	102.49**
		(29.85)			(33.01)	(38.46)	(33.23)
SIZE			-.010		-.022**	-.009	-.021*
			(.010)		(.011)	(.011)	(.011)
K				.001	.019	.070**	.019
				(.032)	(.031)	(.030)	(.031)
Intercept	52.66**	44.93**	56.71**	52.54**	47.15**	30.11**	47.93**
	(4.95)	(6.15)	(6.50)	(6.91)	(8.55)	(9.42)	(8.61)
R-sq	.33	.37	.34	.33	.42	.30	.42

Note: Standard errors in parentheses
(*) indicates significance at the 10% level.
(**) indicates significance at the 5% level.

(UCOV) is the dependent variable.

The weaker coefficients on EXIT than EXIT99 are consistent with the hypothesis that behavior of a fringe of small firms accounting for 1% of employment is not particularly important for unionization, so EXIT is essentially a noisy version of EXIT99 that is subject to attenuation bias. The coefficients on HHI and K have the expected sign, since concentrated and capital-intensive industries have more quasi-rents, and thus pose more attractive targets for unions.

These results are of similar magnitude to those predicted by the model when equation (11) is evaluated using mean values for UMEM and EXIT99 and a range of potential values of λ and $2A(\alpha)B$.[26]

Of course, other models might also suggest a link between unionization and firm lifespan, even after controlling for other variables. For example, unions may have less incentive to invest in organizing short-lived firms. Nevertheless, these results are certainly not inconsistent with this model.

[26] Assuming $U^* = 0.26$, $\lambda = 1.5$, $\delta(0) = .076$, and $2A(\alpha)B = \frac{\delta(0)}{2}$ yields a predicted value (from equation (11) of the coefficient on EXIT99 of 3.8. Changing any of these assumptions by 25% yields predicted coefficients between 2.8 and 5.2.

The link between firm lifespan and unionization may also help explain some other correlates of unionization. Large firms have typically been around longer than small firms, and this may help explain why large firms have higher unionization. (Note that we find no direct effect of firm size on unionization at the industry level, controlling for exit rates, but that we do find a positive, albeit insignificant, effect if we do not control for exit rates, suggesting that the positive correlation between firm size and unionization may be due to lower large firms' exit rates.) The model also suggests that the long lives of public sector institutions may help explain why unionization rates are so high in the public sector and why unions resist privatization. Anecdotal evidence suggests that there is less turnover of firms in Europe. It is conceivable that this could help account for the greater power of unions in Europe than in the United States.

Some observers claim that business is becoming more competitive. This model suggests that if this is the case, and if this is interpreted as in increase in turnover among firms, then it may exacerbate the decline in unionization.

6 Extensions

This section presents two extensions to the model. First, we show that a simple extension of the model suggests that unions that devote more resources to organizing than would be preferred by members will have an evolutionary advantage over others. Next, we show that under a variant of the model, there may be multiple equilibria in unionization and firm turnover. In one equilibrium, unionization is low and turnover among firms is rapid, with only high productivity firms staying in business. In the other equilibrium, unionization is high and turnover among firms is low, with low productivity firms staying in business, rather than being driven out of business by higher productivity new entrants.

6.1 Endogenizing Organizing Effort

We have so far taken B, the amount unions spend per unionized firm on organizing, as exogenous. In fact, unions may also differ in the amount they spend on organizing efforts. As with the determination of α, there are several ways to endogenize the determination of B. The traditional maximizing approach assumes that increased union density increases the union's bargaining power, and asks what level of B would be optimal for members (see Wallerstein [1989]). However, there are certain phenomena that this approach has difficulty explaining. In particular, many unions devote substantial resources to organizing outside their core industries. For example, the Steelworkers organize employees at Chock Full O'Nuts, the Teamsters represent casino workers in Las Vegas, and as discussed above, the UAW organizes graduate students at NYU. While it is possible to see how a steel worker or auto worker might benefit from organizing other workers in their industry, it is harder to see why they would prefer to spend their union dues organizing outside their core industries.

By contrast, our approach takes a worker's preferences over the determination of B as given, and ask what level of B is evolutionarily stable. As in the determination of rent extraction, we argue that there may be a selective advantage to unions that encourage leaders to spend more on organizing efforts than would be optimal for members. As a result, unions controlled by leaders may not only have lower δ but also higher B than would be preferred by members.

A natural way to extend the model to include this case is to include B in the function that indexes how attractive a particular union is to workers, so that $A(\alpha)$ becomes $A(\alpha, B)$. While in many cases the welfare maximizing level of B may be equal to 0, it is conceivable that workers may prefer some organizing expenditure to none. [27] When the relative attractiveness of different levels of α and B are multiplicatively separable, it is easy to show that the evolutionarily stable union has both a lower level of α and a higher level of B than the welfare-maximizing union. When they are not separable, it is possible that the evolutionarily stable union could have either a lower level of α or a higher level of B than the welfare-maximizing union, but not both.

Proposition 7 *Suppose that $A(\alpha, B)$ is multiplicatively separable, i.e. $A(\alpha, B) = A1(\alpha)\, A2(B)$. Then the evolutionarily stable level of α will be less than the welfare maximizing level and the*

[27] Note that the workers' preferences over B are only over the "warm glow" the workers receive from contributing to organizing efforts, rather than, as in Wallerstein [1989], the indirect return from increases in the overall size of the union caused by increased levels of B. The fact that the evolutionary competition will select for higher levels of B that increase union size is an endogenous result of the model, rather than an assumption.

evolutionarily stable level of B will be greater than the welfare maximizing level.

Proof. As in Proposition 3, the evolutionarily stable union is the union with α and B that maximize the ratio $\frac{2A(\alpha,B)B}{\delta(\alpha)}$. Since $A(\alpha, B) = A1(\alpha) A2(B)$, maximizing $\frac{2A(\alpha,B)B}{\delta(\alpha)}$ is equivalent to separately solving two maximization problems:

$$\max_{\alpha} \frac{2A1(\alpha)}{\delta(\alpha)}, \tag{21}$$

$$\max_{B} A2(B)\, B. \tag{22}$$

Since α_W maximizes $A1(\alpha)$, the proof that $\alpha_S < \alpha_W$ is identical to the proof of Proposition 5. To see that $B_W < B_S$, note that $A2(B)$ is maximized at B_W. Since $A2(B) < A2(B_W)$ for all $B < B_W$, $A2(B)\, B < A2(B_W)\, B_W$ for all $B < B_W$. Moreover, by the envelope theorem the derivative of $A2(B)\, B$ at B_W is positive. The value of B that maximizes $A2(B)\, B$ therefore must be greater than B_W. ■

6.2 Multiple Equilibria

As has been demonstrated by Caballero and Hammour [1998] and Moene and Wallerstein [1993], turnover can be lower in a unionized environment. This is because unions will extract a smaller absolute amount from less productive firms that are closer to the exit margin. Unions may thus deter costly entry by new high-productivity entrants without extracting so much from older low-productivity firms that the firms exit. In general equilibrium, therefore, unionization can increase the lifespan of firms. Freeman and Kleiner [1999] provide evidence that unions share rents with firms, extracting less from firms that are in worse economic shape, so that they do not drive the firms into bankruptcy.[28] As discussed in Section 2, a longer firm lifespan encourages the spread of unions. This section uses a variation on the model to show that since high unionization increases firm lifespan, and high firm lifespan increases unionization, there may be one equilibrium with low

[28]Taken literally, these results would imply that union firms do not have a higher death rate than non-union firms, contrary to the assumption in this model. However, Freeman and Kleiner also find that union firms expand less rapidly than non-union firms, and if the unionization rate depends on the existing number of union members, this will have similar effects to the process modeled in this paper. Moreover, it seems possible that Freeman and Kleiner simply cannot detect the effects of unions on the death rates of firms.

unionization and only high productivity firms, and another equilibrium with high unionization in which firms do not exit when productivity falls.

Consider a variant of the model in which there are two productivity levels, H and L, in which output is $Y = H(Q)$ and $Y = L(Q)$ respectively, where Q is employment and $H(Q) > L(Q)$ for all Q. Suppose that $H', L' > 0$ and $H'', L'' < 0$. All new firms have productivity H, but face a Poisson probability δ of switching to productivity L. Whereas in previous sections, we assumed firms exited when they received a negative shock, which can be taken as the case in which $L'(0)$ is sufficiently low, this section considers the case in which $L'(0)$ is sufficiently great that firms may stay in business after a negative shock. Moreover, whereas previous sections examined the case in which δ, the hazard rate of a negative shock, was a function of investment, this section examines the simple case in which δ is exogenous, and hence investment, I, is always zero. Given this, unions set $\alpha = 1$ and extract all rent. (Note, however, that the qualitative results do not depend on these simplifications.) As in Section 2, we continue to assume that $2AB < \delta$, so that if firms exit with any positive hazard rate δ, the unionization rate will be between 0 and 1.

Whereas before we considered a case in which workers had an alternative home production technology, and this tied down the wage, in this section we consider the case in which there is no alternative home production technology (or demand for workers is high enough that all workers are pulled out of the home sector), so that workers are paid their marginal product rather than the reservation wage. Denote the total number of workers as N. We continue to assume that bargaining over employment is statically efficient, so that firms hire workers until the pre-union wage equals the marginal product. (For example, a union contract could allow the firm to hire temporary, non-union workers to raise employment to an efficient level, but then extract the surplus on behalf of members.)

Finally, whereas in the previous section we assumed that there was only one firm for a given product and entry was limited by the need for a license sold at an auction, now we consider the case in which there is a single good produced by multiple firms, and entry is limited by a fixed startup cost. (Think of building a factory, rather than bidding for a cellular telephone license.)

Since there is a single good in the economy, we can normalize the price of this good to 1.[29] Denote the amount of these fixed start-up costs by θ. Free entry ensures that firms enter as long as

$$\theta \leq \mathsf{E}V \tag{23}$$

Note that since N is fixed, when the number of firms F increases, the number of employees per firm, $\frac{N}{F}$, decreases, and because of diminishing returns the marginal product of each employee, and therefore the wage, increases. Since firms' expected profits decline as the pre-union wage increases, firms will enter until $\theta = \mathsf{E}V$. Note that F is therefore no longer fixed, but depends the value of $\mathsf{E}V$, which in turn depends on the chance of being unionized. Denote by F_H the number of firms that will enter in the presence of a union extracting $\alpha = 1$ if all firms are of high productivity. (The precise determination of F_H will be discussed in more detail below.) Note that firms earn zero profits on the marginal worker but earn positive profits on all inframarginal workers, so there is still a flow of annual profits from which they can pay startup costs.

Proposition 8 *If $\theta < \left(\frac{\delta - 2AB}{\delta}\right) \frac{[H(\frac{N}{F_H}) - \frac{N}{F_H} H'(\frac{N}{F_H})]}{\delta + r}$ and $L'(0) < H'(\frac{N}{F_H})$, then there can be two steady-state equilibria: one with partial unionization and only H type firms, and one with complete unionization and only L type firms. Furthermore, if $\theta < \frac{H(n) - nL'(0)}{\delta + r}$, where $n = H'^{-1}(L'(0))$, then multiple equilibria will only occur in the presence of unions.*

Proof. Suppose there are F firms, and all have productivity H, so employment in each is $\frac{N}{F}$. Wages will therefore be $H'\left(\frac{N}{F}\right)$, and pre-union profits will be $\pi(F) = H\left(\frac{N}{F}\right) - \frac{N}{F} H'\left(\frac{N}{F}\right)$. Free entry implies that in steady state, entry costs will equal the probability that the firm remains non-union times the profits of a non-union firm. Assuming that firms that receive the negative shock exit, and that these shocks hit both union and non-union firms with hazard rate δ,

$$p^* = \frac{U^*}{F^*} = \frac{2AB}{\delta}. \tag{24}$$

Recall that we have assumed that $2AB < \delta$, so we know that the union prevalence will be strictly

[29] The results of the model would still go through if there were a downward-sloping demand curve for the good.

between 0 and 1.

Thus, free entry implies that

$$\theta = \left(\frac{\delta - 2AB}{\delta} \right) \left(\frac{H\left(\frac{N}{F}\right) - \frac{N}{F} H'\left(\frac{N}{F}\right)}{\delta + r} \right), \tag{25}$$

where r is the (exogenous) interest rate. There will be a unique equilibrium F_H satisfying this free entry condition, since the assumptions imply that for sufficiently small F, the present discounted value of profits will be greater than θ, and for sufficiently great F, the present discounted value of firms will be less than θ. As discussed above, profits decline monotonically in F, since the greater F, the fewer the workers per firm and the higher wages.

For there to be an equilibrium with only H firms, L firms must not want to produce given wages $H'\left(\frac{N}{F_H}\right)$. This will be the case if $L'(0) < H'(\frac{N}{F_H})$. If this condition is satisfied, firms that receive the negative shock will exit and be replaced by new high-productivity entrants.

In the second steady-state equilibrium, all firms are of type L and unionized. Now, pre-union profits are $\pi(F) = \left[L\left(\frac{N}{F}\right) - \frac{N}{F}L'\left(\frac{N}{F}\right)\right]$ and, since $\alpha = 1$, post-union profits are 0. Since the death rate of L type firms is 0, it is clear from the model in Section 2 that $\frac{U^*}{F^*} = 1$ and the probability that a new firm will be unionized, p^*, is also 1. Since potential entrants know that unions would extract all profits, no new firms enter. Since no firms enter, F is fixed at its initial value, so there can be multiple values of F for which this condition holds.

To see that if $\theta < \frac{H(n) - nL'(0)}{\delta + r}$, where $n = H'^{-1}(L'(0))$, multiple equilibria will only occur in the presence of unions, note that the discounted profits earned by high-type firms exceed startup-costs, no matter how many low-type firms there are. In the absence of a union, high-type firms will therefore always enter if there are only low-type firms present. Therefore, under these conditions, the equilibrium with low productivity firms can only be sustained in the presence of a union. ■

Note that while this exposition has been based on the simplifying assumption that δ is exogenous, it is possible to construct a more general version of the same argument with δ endogenous as in the rest of the paper. The results from such an exercise are essentially similar, though the algebra is substantially more complicated. The main difference in such a model is that α is not necessarily 1. As a result, in the steady-state with low-productivity firms and full unionization, a potential entrant still earns some profits. The condition for there to be a steady-state with lower productivity firms and no entry is that the discounted post-union profits earned by an entrant are lower than the start-up costs, so that entry remains unprofitable.

One could speculate that the U.S. is currently in a low unionization, high firm turnover, high productivity equilibrium, and that Europe has spent much of the last two decades in a high unionization, low turnover, and low productivity equilibrium.

7 Conclusion

This paper has applied techniques from biology to model unions. A key implication of the model is that the unions we observe today are likely to extract less rent than would be optimal for current members, because unions that do so will have a selective advantage over unions that better represent their members' interests. For union leaders to moderate workers' wage demands, however, they must be insulated from workers by incumbency advantages. In fact, these incumbency advantages are widespread among today's unions.

In addition, the model implies that industries with high turnover of firms will have low unionization. Empirical evidence was presented that supports this prediction. Finally, the model also implies that there may be multiple equilibria in unionization levels.

In the conclusion, we discuss the relationship between our model and other theories of incumbency advantages in unions, the normative implications of the analysis, and the applicability of the evolutionary analysis here to other institutions, such as firms.

7.1 Relationship to Other Theories of Incumbency Advantages

The model outlined in this paper is complementary with other, more traditional explanations of incumbency advantages in unions. Sociological explanations, such as Michel's [1949 (1915)] "Iron Law of Oligarchy," suggest that leaders will inevitably seize control of their organizations and work to preserve the organization itself rather than to advance the original goals of the organization. In contrast, the argument here is not that all union leaders will wrest control away from their members due to internal sociological factors and then work to maximize the membership of the union, but rather that those unions that create structures in which this occurs will grow at the expense of unions that narrowly serve their current members' interests. If Michel's process occurs even in a few unions, we will empirically observe these unions much more frequently than unions that are more responsive to their membership.

Another way to explain the typically more moderate position of union leadership is through models in which union leaders are agents whose interests differ from those of their principals, the rank and file.[30] As discussed above, union leaders may plausibly prefer more moderate policies than their members, for example if they are motivated by the empire-building motives often assumed in corporate finance. They may also collude with firm management to moderate wage demands in exchange for contract provisions benefiting the union and its leaders. These considerations may well be the proximate cause of moderation of wage policy by union leaders. However, standard agency theory implies that principals should design optimal mechanisms for agents. It thus begs the question of why so many unions have constitutional institutions that exacerbate agency problems in controlling leaders, such as indirect elections, secret lists of locals and members, and no prohibitions on campaign donations from union staff. In contrast, this biological model suggests that unions with constitutional procedures that exacerbate agency problems will outcompete others that do not.

7.2 Normative Implications

The normative implications of the analysis are ambiguous. Rent extraction by unions reduces firms' expenditure on startup costs and lowers the level of investment chosen by unionized firms. This increases the death rate of unionized firms, which means that the startup costs, though lower than without unions, are paid more frequently.

The welfare implications of these changes depend on the interpretation of investment and start-up costs. Investment and start-up costs may be productive, such as investment in research and

[30] We thank Jean Tirole for pointing this out.

development of improved products, or unproductive, such as advertising designed to establish market leadership for a dot.com seeking first-mover advantage. If up-front investments are productive, as with research and development, having more firms perform this investment may be socially beneficial. If expenditures are unproductive, rent extraction by unions, and therefore lower start-up costs, may potentially represent a benefit to society.

As shown above, the evolutionarily stable level of rent extraction will lead to more unionization in the steady-state than the welfare-maximizing level of rent extraction, but the overall effect on total rent extracted by unions is ambiguous. Because the level of rent extracted from each firm is lower, there will be more ongoing investment by unionized firms than under the welfare-maximizing rent extraction level. The effect on startup-cost expenditure, however, is ambiguous. On the one hand, since the steady-state chance of a new firm being unionized, p^*, is maximized by the evolutionarily stable union, the chance of a new firm being unionized is higher, reducing the expenditure on start-up costs. However, the cost of being unionized, $\alpha\pi$, is lower, increasing the *ex ante* value of the firm. Furthermore, the death rate of firms will be lower, so startup costs will be paid less frequently. The total effect on startup costs is unclear. The general-equilibrium welfare effects of having a lower level of rent extraction are thus ambiguous.

Regardless of these general equilibrium effects, however, the model implies that unions are not extracting the optimal level of rent for their workers. Changing union constitutions to reduce incumbency advantages will likely lead to increased welfare for the union's current members, though it will also reduce long-term unionization.

7.3 Applicability to Other Organizations

Similar evolutionary arguments could be made about organizations other than unions. For example, those religions that grow may be those that are most successful at retaining members, rather than those that maximize members' welfare. Universities whose boards accumulate large endowments may be more likely to survive than universities that pay out from the endowment less conservatively, whether or not this contributes to the universities' educational and research mission. As Dutta and Radner (1999) suggest, firms that maximize their stockholders' interests by paying out dividends may eventually be outnumbered by firms that retain earnings as a safety net, because paying out dividends makes firms more vulnerable to negative shocks.

Reality is likely to lie between the predictions of models in which institutions maximize their owners' welfare and biological models in which organizational characteristics are fixed. The more that members have opportunities to control their organizations, the closer reality is likely to lie to the welfare-maximizing model. For example, the model presented in this paper suggests that if some unions are taken over by their leaders, these unions will displace member-controlled unions. One could consider a more complicated model in which member-controlled unions face a hazard rate of capture by leaders and leader-controlled unions face a hazard rate of reverting to the control of members. In this case, there will be a mixture of leader- and member-controlled unions in steady state. The proportion of each type of union will depend on the hazard rates of transition of control and also on the speed with which member-controlled unions are displaced by unions with incumbency advantages in the evolutionary competition emphasized in this model. If control transitions are rare relative to this speed of displacement, the biological model will be a fairly good predictor of union behavior. If unions with incumbency advantages take a long time to displace those that serve their members perfectly, the economic model of welfare-maximizing unions will describe a greater proportion of unions.

This suggests that firms may be closer to the welfare-maximizing end of the spectrum than unions, since control of unions by members is likely to be weaker than control of firms by shareholders. There is a substantial free-rider problem for workers in controlling union management, just as there is an important free-rider problem for shareholders in controlling firm management. However, in many cases, firms will have one large shareholder with a substantial stake in firm governance. In contrast, no single union member has a substantial stake in reforming the union leadership. Moreover, whereas there is a large financial incentive for outsiders to take over firms managed against shareholders' interests, there is much less incentive for outsiders to challenge existing unions for the right to represent workers.

A Appendix

The first part of the appendix gives some of the proofs omitted from the main text. The second part discusses the behavior of the model outside of the steady-state. The third part shows the instability of the trivial steady state with $U^* = 0$ and the stability of the steady-state with $U^* = \frac{2\delta(0)A(\alpha)B}{\delta(\alpha)^2 - 2A(\alpha)B[\delta(\alpha)-\delta(0)]}$. The final part of the appendix presents a simple proof that in the bargaining game between union leaders and the rank-and-file, union leaders obtain their most preferred outcome.

A.1 Proofs

Proof of Lemma 1. For clarity of exposition, this proof will consider the case of a steady-state with two unions. However, the same arguments go through in the cases when there are more than two unions in the steady-state. As will be shown, however, there can be more than two unions in the steady state only if there is some value q such that there are more than two distinct levels α such that $\frac{2A(\alpha)B}{\delta(\alpha)} = q$, which will only occur under parameterizations of $A(\alpha)$ and $\delta(\alpha)$ such that $\frac{2A(\alpha)B}{\delta(\alpha)}$ has more than 1 critical point.

Recall that in the steady-state in which a single union has organized all firms with difficulty levels less than or equal to p, equation (5) stated that a union must spend its entire organizing budget organizing new firms with difficulty levels less than or equal to p. Adapting this condition to the case of two unions yields

$$A(\alpha_M)BM = [\delta(\alpha_M)M + \delta(\alpha_R)R + \delta(0)(1-U)] \frac{A(\alpha_M)BM}{A(\alpha_M)BM + A(\alpha_R)BR} \frac{p_M^2}{2}. \qquad (26)$$

If p_M and p_R were different, then this equation would apply only to the union with the smaller p. Supposing for the moment that M had the lower p (though in practice it could be either M or R), then the union R would be able to organize all unions in the interval $[p_M, p_R]$ instead of just the fraction $\frac{A(\alpha_R)BR}{A(\alpha_M)BM+A(\alpha_R)BR}$ of them. However, rewriting equation (26) shows that

$$p_M = \sqrt{\frac{2\left[A\left(\alpha_M\right)B + A\left(\alpha_R\right)B\right]}{\delta\left(\alpha_M\right)M + \delta\left(\alpha_R\right)R + \delta\left(0\right)\left(1 - M - R\right)}}. \tag{27}$$

Inspection of equation (27) shows that p_M and p_R must be the same for both unions in the steady-state since the equation for p_R would be exactly the same. Therefore we know that in the steady state the set of firms being organized each period by both unions have the same difficulty profile. This, in turn, is a consequence of allocating firms in proportion to the unions' effective organizing budget.

The second condition for the steady state is that $\dot{U} = 0$, so that the size of the union remains the same. Since the union's entire budget is exhausted in organizing newly created firms, in the steady state we know that, for $\dot{U} = 0$,

$$\delta\left(\alpha_M\right)M = \left[\delta\left(\alpha_M\right)M + \delta\left(\alpha_R\right)R + \delta\left(0\right)\left(1 - M - R\right)\right]\frac{A\left(\alpha_M\right)BM}{A\left(\alpha_M\right)BM + A\left(\alpha_R\right)BR}p_M \tag{28}$$

and the equivalent equation for R. This equation states that the number of member firms lost due to negative shocks must be exactly replaced by the number of firms organized during the same period. There are $\left[\delta\left(\alpha_M\right)M + \delta\left(\alpha_R\right)R + \delta\left(0\right)\left(1 - M - R\right)\right]$ firms created each period, of which the M union targets the fraction $\frac{A(\alpha_M)BM}{A(\alpha_M)BM+A(\alpha_R)BR}$ and from which it organizes all firms with difficulty levels below p_M. Substituting equation (27) for p_M yields the steady-state condition

$$\frac{A\left(\alpha_M\right)B}{\delta\left(\alpha\right)} = \sqrt{\frac{A\left(\alpha_M\right)BM + A\left(\alpha_R\right)BR}{2\left[\delta\left(\alpha_M\right)M + \delta\left(\alpha_R\right)R + \delta\left(0\right)\left(1 - M - R\right)\right]}}. \tag{29}$$

By substituting equation (29) into equation (27), we can see that

$$\frac{2A\left(\alpha_M\right)B}{\delta\left(\alpha_M\right)} = p_M = p_R = \frac{2A\left(\alpha_R\right)B}{\delta\left(\alpha_R\right)}. \tag{30}$$

The algebra would have been essentially similar if there had been more than two types of union. ∎

Proof of Proposition 4. Suppose that the steady state contains an incumbent union, α_I.

Lemma 1 guarantees that if there are additional unions in the steady state with different α, those unions will have the same value of $p^* = p_I^*$. Therefore, in the steady-state, all firms with difficulty level less than p_I^* will be unionized and all firms with higher difficulty levels will not be unionized. For simplicity, the remainder of the proof focuses on the case where there is only one union in the steady state, but because the ratio $\frac{2A(\alpha)B}{\delta(\alpha)}$ is the same for all incumbent unions in a steady state, the same arguments go through when there are multiple incumbent unions.

Consider an invasion by a union that extracts α_S with size $S < \varepsilon$, where ε is very close to 0. Using a similar argument to the one in Proposition 3, we can see that $p_S \approx p_I^*$. Therefore, the initial growth of the union will be approximately

$$\dot{S} \approx \left[\delta\left(\alpha_I\right) I + \delta\left(\alpha S\right) S + \delta\left(0\right)\left(1 - I - S\right)\right] \frac{A\left(\alpha_S\right) BS}{A\left(\alpha_I\right) BI + A\left(\alpha_S\right) BS} p_I^* - \delta\left(\alpha_S\right) S. \tag{31}$$

To see that this growth is positive, observe that S is approximately equal to 0 and recall that in the steady state, $p_I^* = \frac{2A(\alpha_I)B}{\delta(\alpha_I)} = \sqrt{\frac{2A(\alpha_I)BI}{\delta(\alpha_I)I+\delta(0)(1-I)}}$. This allows us to simplify this expression and write

$$\frac{\dot{S}}{S} \approx \frac{2A\left(\alpha_S\right) B}{p_I^*} - \delta\left(\alpha_S\right) \tag{32}$$

which is greater than 0 since $\frac{2A(\alpha_S)B}{\delta(\alpha_S)} > \frac{2A(\alpha_I)B}{\delta(\alpha_I)}$. This means that the invading union will grow. ■

Proof of Proposition 6. Recall from equation (6) that in the steady state, the number of unionized firms that die each instant must exactly match the number of newly created forms unionized in that instant. Rearranging equation (6) yields the condition

$$\frac{2A\left(\alpha\right) B\delta\left(0\right)}{\delta\left(\alpha\right)} = U^* \left(\delta\left(\alpha\right) - \frac{2A\left(\alpha\right) B}{\delta\left(\alpha\right)} \left[\delta\left(\alpha\right) - \delta\left(0\right)\right]\right). \tag{33}$$

Consider a change from α to α_S. Denote by ΔP the difference in the fraction of firms unionized each period, i.e. $\Delta P = \frac{2A(\alpha_S)B}{\delta(\alpha_S)} - \frac{2A(\alpha)B}{\delta(\alpha)}$, and denote by $\Delta\delta$ the same change in $\delta\left(\alpha\right)$, i.e. $\Delta\delta = \delta\left(\alpha_S\right) - \delta\left(\alpha\right)$. Since α_S maximizes the ratio $\frac{2A(\alpha)B}{\delta(\alpha)}$, ΔP will be greater than 0, and since $\alpha_S < \alpha$ by assumption, $\Delta\delta$ will be less than 0.

Since the left hand side of equation (33) is higher under α_S than under α, U^* will be higher under α_S than under α if the right hand side is lower. The change in the right hand side between α and α_S will be

$$\Delta\delta\left[1-\frac{2A\left(\alpha\right)B}{\delta\left(\alpha\right)}\right]-\left[\delta\left(\alpha_S\right)-\delta\left(0\right)\right]\Delta P. \tag{34}$$

>From condition (9), we know that $\frac{2A(\alpha)B}{\delta(\alpha)}<1$ for all α, and since $\delta\left(\alpha\right)\geq\delta\left(0\right)$ for all α, we know that $\delta\left(\alpha_S\right)-\delta\left(0\right)\geq 0$. Therefore the expression for the change in the right hand side, equation (34), will be less than 0. We can therefore conclude that $U^*\left(\alpha_S\right)>U^*\left(\alpha\right)$. ■

A.2 Dynamics

Outside of the steady-state, the state-space can be characterized by the number of union firms, U, and the difficulty distribution of all unorganized firms. As discussed above, in the steady-state the distribution of non-unionized firms' difficulties is simply uniform from the threshold p to 1, but in certain kinds of transitions–for example, those in which the difficulty level below which all firms are unionized, p, is shrinking–the distribution can be non-uniform. To track the dynamics, then, one needs keep track not only of the transition equations for U and p, but also the transition equation for the entire difficulty distribution. These transition equations are used in Section 3 to characterize the evolutionarily stable steady-state.

At any instant, assuming that there is no discontinuous increase in the number of firms, there are two different sets of firms that the union may chose to organize: the "thick" set of firms that are non-unionized and the "thin" set of firms that were created that instant to replace firms that exited due to a negative shock. The number of non-unionized firms is in the thick set is $1-U$ and the number of firms in the thin segment is

$$\left[\delta\left(\alpha\right)U+\delta\left(0\right)\left(1-U\right)\right]dt \tag{35}$$

Facing this profile of non-unionized firms, the union will organize the easiest firms it can. These will be all of the firms in the thin segment with cost less than p and then as many firms in the thick segment as it can with whatever remains of its organizing budget at that moment. Note that p represents the lower bound of the "thick" set of non-unionized firms–it will be possible in certain transitions that there are unionized firms whose difficulties are greater than p. Since the distribution of firms in the thin segment is uniform, the cost of organizing all firms in the thin segment with cost less than p will be

$$[\delta(\alpha)U + \delta(0)(1-U)]\, dt\frac{p^2}{2} \tag{36}$$

so that the budget surplus or effective deficit becomes

$$A(\alpha)BU - [\delta(\alpha)U + \delta(0)(1-U)]\, dt\frac{p^2}{2} \tag{37}$$

If the budget has a surplus, then the growth of the union will be the number of firms in the thin segment with difficulty levels less than or equal to p plus however many older firms the union can afford to organize at marginal cost p with whatever remains of its budget, minus the number of its member firms it lost due to negative shocks:

$$\dot{U} = [\delta(\alpha)U + \delta(0)(1-U)]\, p + \frac{A(\alpha)BU - [\delta(\alpha)U + \delta(0)(1-U)]\,\frac{p^2}{2}}{p} - \delta(\alpha)U \tag{38}$$

On the other hand, if the union's budget is not sufficient to organize all firms in the thin segment with costs less than or equal to p, the union will organize as many of those firms as it can. This will be all newly created firms with difficulty levels less than or equal to some cutoff level l such that the total budget exactly equals the cost of organizing the firms, i.e.

$$[A(\alpha)BU]\, dt = [\delta(\alpha)U + \delta(0)(1-U)]\, dt\frac{l^2}{2} \tag{39}$$

This implies that

$$l = \sqrt{\frac{2A(\alpha)BU}{[\delta(\alpha)U + \delta(0)(1-U)]}} \tag{40}$$

The change in the number of unionized firms in this case will therefore be the fraction l of thin firms unionized, multiplied by the total number of thin firms, less the number of unionized firms that exit:

$$\dot{U} = \sqrt{2A(\alpha)BU\left[\delta\left(\alpha\right)U + \delta\left(0\right)\left(1-U\right)\right]} - \delta\left(\alpha\right)U \tag{41}$$

Keeping track of changes in the distribution of the non-unionized firms is somewhat trickier. Suppose that the density of non-union firms in the thick segment at some difficulty level c is $f\left(c\right)$. To find $\dot{f}\left(c\right)$ for those levels c that remain non-unionized (which will be all $c \geq p$) it will be instructive to consider the discrete case and take limits, so suppose that the density is the same over some small segment dz and small amount of time dt. Denote by $f_0\left(c\right)$ the density of firms in the segment dz before the time starts and $f_1\left(c\right)$ the density after the unit of time has passed. Define the density so that the total number of firms in the segment dz before the change will be $\left(1-U\right)f_0\left(c\right)dz$ and after the change will be $\left(1-U-\dot{U}dt\right)f_1\left(c\right)dz$. The number of firms after the change will be equal to the number of firms in the segment before the change plus the number of firms that are born with costs in the segment minus the number of firms in the segment that exit due to the shock:

$$\begin{aligned}\left(1-U-\dot{U}dt\right)f_1\left(c\right)dz = f_0\left(c\right)\left(1-U\right)dz + \\ \left[\delta\left(\alpha\right)U + \delta\left(0\right)\left(1-U\right)\right]dtdz - \delta\left(0\right)\left(1-U\right)f_0\left(c\right)dtdz\end{aligned} \tag{42}$$

The change in f will therefore be

$$\begin{aligned}\dot{f}dt &= \left[f_1\left(c\right)dz - f_0\left(c\right)dz\right]dt \\ &= \frac{f_0\left(c\right)\left(1-U\right)dz + \left[\delta\left(\alpha\right)U + \delta\left(0\right)\left(1-U\right)\right]dtdz - \delta\left(0\right)f_0\left(c\right)\left(1-U\right)dtdz}{\left(1-U-\dot{U}dt\right)} - f_0\left(c\right)dz\end{aligned} \tag{43}$$

Simplifying and taking limits yields the equation for $\dot{f}$:

$$\dot{f}\left(c\right) = \frac{\left[\delta\left(\alpha\right)U + \delta\left(0\right)\left(1-U\right)\right] - \left[\delta\left(0\right)\left(1-U\right) - \dot{U}\right]f_0\left(c\right)}{\left(1-U\right)} \tag{44}$$

Note that substituting in the steady-state value of U and setting $\dot{f}\left(c\right)$ and $\dot{U}$ equal to 0 yields a steady-state value for $f_0(c)$ of $\frac{1}{1-p^*}$, which means that the distribution of costs of non-unionized firms in the steady-state is uniform over the range $[p, 1]$, as expected.

We also need to keep track of changes to p, the lower bound of the support set of the thick segment. If the union has a budget surplus (i.e. equation (37) is positive), the union has organizing funds remaining after unionizing all firms in the thin segment with costs less than or equal to p. The change in p will therefore be equal to the number of new firms unionized at cost p divided by the density of firms at that cost level, i.e.

$$\dot{p} = \frac{A(\alpha)BU - [\delta(\alpha)U + \delta(0)(1-U)]\frac{p^2}{2}}{p} \frac{1}{(1-U)f(c)}. \tag{45}$$

On the other hand, when the union's organizing budget is not sufficient to unionize all newly created firms with costs less than or equal to p, the new value of p will be the highest-cost firm that the union is able to unionize, i.e.

$$p = \sqrt{\frac{2A(\alpha)BU}{[\delta(\alpha)U + \delta(0)(1-U)]}}. \tag{46}$$

Together, the transition equations $\dot{U}$, $\dot{f}(c)$, and $\dot{p}$ completely characterize the dynamics of the system.

A.3 Stability

As discussed in Section 2.2, the model has two steady-states, the trivial steady-state where $U^* = 0$ and the non-trivial steady state where $U^* = \frac{2\delta(0)A(\alpha)B}{\delta(\alpha)^2 - 2A(\alpha)B[\delta(\alpha)-\delta(0)]}$. This section gives proofs that the first steady state is locally unstable and the second is locally stable.

Proposition 9 *The steady-state with $U^* = 0$ is locally unstable.*

Proof. Consider starting out from the steady-state of $U = 0$ and introducing a union of size $\varepsilon > 0$. Assume that this union consists of the least-costly ε firms, so that the remaining non-unionized firms have costs uniformly distributed on the interval $[\varepsilon, 1]$. This assumption makes it the hardest to show instability, because the cost distribution facing the union is the highest possible.

Recall from equation (37) that the budget surplus or deficit will be given by

$$A(\alpha)B\varepsilon - [\delta(\alpha)\varepsilon + \delta(0)(1-\varepsilon)]\,dt\frac{\varepsilon^2}{2} \tag{47}$$

since p will be equal to ε. Note that the average organizing costs faced by the union will be less than ε since it will be organizing some newly created firms in the thin segment $[0,\varepsilon]$ and some in the thick segment at ε. The growth rate of the union will therefore be greater than it would be if it spent its entire budget organizing firms with cost ε, i.e.

$$\dot{U} > \frac{A(\alpha)B\varepsilon}{\varepsilon} - \delta(\alpha)\varepsilon \tag{48}$$

which will be clearly positive for ε small enough. ■

Proposition 10 *The steady-state with* $U^* = \frac{2\delta(0)A(\alpha)B}{\delta(\alpha)^2 - 2A(\alpha)B[\delta(\alpha)-\delta(0)]}$ *is locally stable.*

Proof. Consider first a union in the steady state where ε of the firms in the union revert back to non-union status. The union's organizing budget will therefore be $A(\alpha)B(U-\varepsilon)$. Denote by f the highest cost level firm in the thin segment the union could organize with such a budget, and by p' the lowest cost value of firms in the thick segment. We know that $p' \leq p$, but the precise value will depend on the cost level of the ε firms that switched from being unionized to being non-unionized. If $f > p'$, the union will spend the remaining budget surplus organizing the thick segment of firms with cost p'; otherwise it will organize as many firms in the thin segment as it can. The growth of the union will therefore be greater than or equal to the growth if it spend its entire organizing budget on firms with costs less than or equal to f, i.e.

$$\dot{U} \geq \sqrt{2A(\alpha)BU^*[\delta(\alpha)U^* + \delta(0)(1-U^*)]} - \delta(\alpha)U^* \tag{49}$$

Substituting in $U^* - \varepsilon$ for U and rearranging terms, we can see that, for $\varepsilon > 0$, the growth will be positive (and therefore the steady-state will be stable) if

$$2A(\alpha)B[\delta(\alpha)(U^*-\varepsilon) + \delta(0)(1-U^*+\varepsilon)] > \delta^2(\alpha)(U^*-\varepsilon) \tag{50}$$

Since $\dot{U} = 0$ at U^*, the U^* terms in this expression cancel, and we are left with the condition

$$2A(\alpha)B\left[\delta\left(\alpha\right)\varepsilon - \delta\left(0\right)\varepsilon\right] < \delta^2\left(\alpha\right)\varepsilon \tag{51}$$

Since condition (9) guarantees that $2A\left(\alpha\right)B \leq \delta\left(\alpha\right)$, this condition will be satisfied and the union will return to the steady-state.

Next, consider a union in the steady state where ε of the non-union firms spontaneously unionize. To make it hardest to show stability, assume that these firms are the costliest to unionize, i.e. the firms with costs from the interval $\left[1 - \frac{\varepsilon(1-p)}{(1-U)}, 1\right]$. This is the most difficult assumption for showing stability since we have removed the firms that are costliest to organize. The first thing to check is whether the union will have sufficient organizing funds left over to begin organizing the thick segment, i.e. whether or not equation (37) is positive. The budget surplus will be

$$A\left(\alpha\right)B\left(U^* + \varepsilon\right) - \left[\delta\left(\alpha\right)\left(U^* + \varepsilon\right) + \delta\left(0\right)\left(1 - U^* - \varepsilon\right)\right]\frac{p^2}{2} \tag{52}$$

Since $\dot{U}^* = 0$ at the steady state, the terms from equation (5) cancel, so the surplus will be greater than or equal to 0 if

$$A\left(\alpha\right)B\varepsilon - \left[\delta\left(\alpha\right)\varepsilon - \delta\left(0\right)\varepsilon\right]\frac{p^2}{2} \geq 0 \tag{53}$$

Since p is equal to $\frac{2A(\alpha)B}{\delta(\alpha)}$ in the steady-state, we know that this condition will hold if

$$\delta\left(\alpha\right)^2 \geq 2A\left(\alpha\right)B\varepsilon\left[\delta\left(\alpha\right)\varepsilon - \delta\left(0\right)\varepsilon\right] \tag{54}$$

which is exactly the same as inequality (51), and holds by the same logic.

Given that there is a budget surplus, the change in U will be given by

$$\dot{U} = \left[\delta\left(\alpha\right)\left(U^* + \varepsilon\right) + \delta\left(0\right)\left(1 - U^* - \varepsilon\right)\right]p + \frac{A(\alpha)B\left(U^* + \varepsilon\right) - \left[\delta\left(\alpha\right)\left(U^* + \varepsilon\right) + \delta\left(0\right)\left(1 - U^* - \varepsilon\right)\right]\frac{p^2}{2}}{p} - \delta\left(\alpha\right)\left(U^* + \varepsilon\right) \tag{55}$$

Once again, since $\dot{U}^* = 0$ at the steady state, canceling out the terms from equation (5) and (6) yields

$$\dot{U} = [\delta(\alpha)\varepsilon - \delta(0)\varepsilon]\, p + \frac{A(\alpha)B\varepsilon - [\delta(\alpha)\varepsilon - \delta(0)\varepsilon]\frac{p^2}{2}}{p} - \delta(\alpha)\varepsilon \quad (56)$$

Substituting in for p and rearranging terms yields inequality (51) as the condition for $\dot{U} < 0$. Since we have already shown that this inequality holds, the growth rate of the union will be negative and it will return to the steady-state. ■

A.4 Approval of Union Contracts[31]

The need for the rank-and-file to approve a contract is similar to a repeated bargaining game, where the union leaders (the "proposer") make take-it-or-leave it offers to the rank-and-file (the "acceptor"). Each time the rank-and-file reject the proposal, both sides lose a percentage γ of the total amount they are bargaining over due to costs of delay. To show that the proposer has complete bargaining power in such a game, we will consider a game involving bargaining over a dollar. This is analogous to bargaining over different levels of utility received from different levels of α. The following proposition shows that the unique Subgame-Perfect Nash Equilibrium of this game is that the union leaders receive their ideal level of α, even though the rank-and-file must approve all contracts.

Proposition 11 *In the repeated take-it-or-leave it bargaining game over a fixed amount with $\gamma < 1$, the unique Subgame-Perfect Nash Equilibrium is for the proposer to obtain the full amount being bargained over.*

Proof. Suppose that the two parties are bargaining over how to split one dollar. Let M be the supremum of the set of expected payoffs that the acceptor can get in any Sub-Game Perfect Equilibrium of the game beginning with an offer. Note that since every subgame that begins with an offer has the proposer making an offer, all such subgames admit the same equilibrium.

In any subgame, the acceptor must accept any offer greater than or equal to γM, since if the offer is rejected the most that the acceptor can hope to receive is γM. But this means that M is

[31] We thank Keith Chen for this proof.

weakly less than γM. This can only be the case if $M = 0$. Thus, $M = 0$ and the proposer obtains the full amount. ■

References

[1] Anderson, Roy M., and May, Robert M., *Infectious Diseases of Humans: Dynamics and Control*. Oxford University Press [1991].

[2] Benson, Herman, [1986]. "The Fight for Union Democracy," in *Unions in Transition: Entering the Second Century*, Seymour Martin Lipset, ed. San Francisco: Institute for Contemporary Studies, pp. 323-373.

[3] Caballero, Ricardo J. and Hammour, Mohamad L, "The Macroeconomics of Specificity," *Journal of Political Economy* 106(4), [August, 724-767, 1998].

[4] Dickens, William T. and Jonathan S. Leonard, "Accounting for the Decline in Union Membership, [1950-1980]," *Industrial and Labor Relations Review* 38(3), [April, 323-334, 1985].

[5] Dunne, Timothy; Roberts, Mark J; Samuelson, Larry, "Patterns of Firm Entry and Exit in U.S. Manufacturing Industries," *Rand Journal of Economics* 19 (4), [Winter, 495-515, 1988].

[6] Dutta, Prajit and Radner, Roy, "Profit Maximization and the Market Selection Hypothesis," *Review of Economic Studies* 66, [October, 769-798, 1999].

[7] Ewald, Paul W. *Evolution of Infectious Disease.* Oxford and New York: Oxford University Press, 1994.

[8] Moene, Karl Ove, Wallerstein, Michael, and Michael Hoel. "Bargaining Structure and Economic Performance. In eds. Flanagan, R. J., Moene, K. O., and M. Wallerstein, *Trade Union Behavior, Pay-Bargaining, and Economic Performance.* FIEF Studies in Labour Markets and Economic Policy. Oxford; New York; Toronto and Melbourne: Oxford University Press, Clarendon Press, 1993.

[9] Freeman, Richard B., "Spurts In Union Growth: Defining Moments And Social Processes." In eds. Bordo, Michael D; Goldin, Claudia; and Eugene N. White, *The Defining Moment: The Great Depression and the American Economy in the Twentieth Century.* NBER Project Report Series. Chicago and London: University of Chicago Press, 1998.

[10] Clark, Kim and Richard B. Freeman, "How Elastic is the Demand for Labor? A Reply." *Review of Economics and Statistics* 65(4) [November, 694, 1983].

[11] Freeman, Richard B. and Morris M. Kleiner, "Do Unions Make Enterprises Insolvent?" *Industrial and Labor Relations Review* 52 (4) [July, 510-27, 1999].

[12] Geoghegan, Thomas, [1992]. *Which side are you on? Trying To be For Labor When It's Flat On Its Back*. New York: Farrar, Straus, & Giroux.

[13] Hannan, M.T., and J. Freeman. "The Ecology of Organizational Founding: American Labor Unions, [1836-1985]." *American Journal of Sociology* 92: 910-43 [1987].

[14] Hannan, M. T., and J. Freeman, The Ecology of Organizational Mortality: American Labor Unions, 1836-1985. American Journal of Sociology 94: 22-52 [1988].

[15] Hirsch, Barry T; MacPherson, David A. "Union Membership and Coverage Files from the Current Population Surveys: Note." *Industrial & Labor Relations Review* 46 (3), [April, 574-78, 1993].

[16] Kleiner, Morris M., and Pilarski, Adam M. "Does Internal Union Political Competition Enhance Its Effectiveness?" In Samuel Estreicher, Harry Katz, and Bruce Kaufman, eds., *Internal Governance and Organizational Effectiveness of Labor Unions*, New York: Kluwer Publishing, forthcoming.

[17] Lipset, Seymour Martin, Martin Trow, and James S. Coleman, *Unions Democracy*. NY: Free Press [1956].

[18] Michels, Robert. 1949. *Political Parties*. Translated by Edward Cedar Paul. Glencoe, IL: Free Press [1915].

[19] Nelson, R.P., and S.G. Winter, *An Evolutionary Theory of Economic Change*. Cambridge, MA: Harvard University Press [1982].

[20] Ross, Arthur M., *Trade Union Wage Policy*. University of California Press. Berkeley and LA [1950].

[21] Tornell, Aaron, "Rational Atrophy: the U.S. Steel Industry." NBER Working Paper Series #6804 [1997].

[22] Wallerstein, Michael, "Union Organization in Advanced Industrial Democracies." *American Political Science Review* 83 (2), [June, 481-501, 1989].

[23] Weiler, Paul. "Striking a New Balance: Freedom of Contract and the Prospects for Union Organization." *Harvard Law Review* 98(351), [December, 1984].