

A Biological Model of Unions[†]

By MICHAEL KREMER AND BENJAMIN A. OLKEN*

This paper applies principles from evolutionary biology to the study of unions. We show that unions that implement the preferred wage and organizing policies of workers will be displaced in evolutionary competition by unions that either extract less from firms, allowing them to live longer, or spend more on union organizing, or both. This implies that unions with constitutional incumbency advantages that allow leaders to depart from members' preferences may have a selective advantage, allowing them to grow at the expense of unions lacking such provisions. Evidence from the history of American unions supports these predictions. (JEL A12, J51)

This paper applies concepts from evolutionary biology to the interaction between unions and firms. We argue unions that implement the wage and organizing policies preferred by workers will be displaced in competition with unions that either extract less rent from firms, allowing them to live longer, or spend more on union organizing, allowing them to organize more new workers.

In particular, we show that a union that implements workers' preferences will not be evolutionarily stable. It can be outcompeted by a union that either reduces the level of rents obtained for workers from the level that would balance workers' concern for higher wages with their concern for protecting their jobs, or increases organizing expenditures beyond the level that maximizes current union members' welfare. Were unions to take actions that maximize the welfare of existing members, they would not be taking into account externalities on potential future union members.

In biological models, whether a trait spreads depends not on how well it serves the welfare of the organism, but rather on its impact on reproduction. Biological models suggest that selection pressure often works against organisms that are too harmful to their hosts. For example, an organism such as the Ebola virus, which kills its host in days and is transmissible only by direct contact, has little opportunity to spread from one host to another. By contrast, the viruses that cause the common cold

* Kremer: Harvard University, Littauer M-20, Harvard University, Cambridge, MA 02138, Brookings Institution, Center for Global Development, and National Bureau of Economic Research (e-mail: mkremer@fas.harvard.edu); Olken: Massachusetts Institute of Technology, 50 Memorial Drive, Cambridge, MA 02142, and National Bureau of Economic Research (e-mail: bolken@mit.edu). We are very grateful to Radu Ban, Kaushal Challa, Marcos Chamon, Rohit Gupta, Sean Flynn, William Hawkins, Jessica Leino, David Mericle, Dan Wood, and Alexander Veytsman for excellent research assistance; to Abhijit Banerjee, Ricardo Caballero, John Dunlop, Susan Dynarski, Henry Farber, Robert Gibbons, Lawrence Mishel, Kevin M. Murphy, Michael Piore, Andrei Sarychev, Jean Tirole, Michael Wallerstein, and three anonymous referees for helpful comments; and to Richard Freeman for the many conversations that inspired the paper. None of them should be held responsible for the results.

[†] To comment on this article in the online discussion forum visit the articles page at: <http://www.aeaweb.org/articles.php?doi=10.1257/app.1.2.150>.

are widespread. Mitochondria, which were probably originally parasites, evolved to become essential to their hosts and are now universal (Philip John and F. R. Whatley 1975). Similarly, a trait that reduces individual survival, but increases survival of genetically related organisms, can spread in a population. Animals often take actions that reduce their own expected survival time but increase the expected number of descendants. By analogy, we show that unions that extract less from their firms and/or spend more on organizing will tend to survive more years, and will therefore be able to organize more new unionized firms, thus generating more total “offspring” and ultimately outcompeting other unions in the population.

In biology, evolutionary pressures may be the ultimate cause of an organism’s behavior, but these pressures must work through the proximate cause of specific genes and the chemical processes they induce. While evolutionary pressures may select for particular union wage and organizing policies, union members will continue to choose policies and leaders that maximize their welfare in the absence of specific institutions that lead the union to behave otherwise. Sustaining a different policy therefore requires constitutional institutions—the analogue of an organism’s genes—that allow union leaders to systematically deviate from the policies preferred by workers without being removed from office. We argue that institutions that create incumbency advantages for union leaders can serve precisely this function, and hence selection pressure favors unions with such institutions over those with weaker incumbency advantages.

It is worth being explicit about how our argument differs and does not differ from the existing literature. It has long been recognized that if unions extract too much from firms they will put firms out of business. Our contribution is not this, but rather showing that the behavior of the union that maximizes welfare for its members will differ from that of an evolutionarily stable union, due to the externalities union organizing expenditure creates on organizing future union members. This implies that unions with constitutions that incorporate incumbency advantages for leaders could potentially have evolutionary advantages over unions that are more responsive to the wishes of existing members.

There is evidence for this implication. In particular, we empirically show that unions with indirect leadership elections in 1955 were less likely to decline over the subsequent 40 years than those unions with direct elections for union leaders. More broadly, we document how over the history of the US union movement, from the Knights of Labor in the 1880s to the creation of the American Federation of Labor (AFL) and the Congress of Industrial Organizations (CIO), there has been a general trend away from independent, local unions and toward national unions with more centralized authority and less direct control by the rank and file, as would be predicted by the model. The recent split in the AFL-CIO can be seen in light of the model as an attempt by some of the unions with more central control to increase organizing expenditures.

The model is consistent with the observation that rank and file dissident movements almost always demand that unions deliver more rents to workers than the incumbent union leadership demanded, suggesting that incumbency advantages lead to union policies that extract lower wages than would be optimal for workers. In several cases in which incumbency advantages have been weakened due to plausibly exogenous federal interventions, dissident movements have become powerful and union membership declined.

Our model differs from a simple supply and demand analysis of unionization because it suggests that features that might seem to make a union less attractive to workers by causing the union to act in ways that do not maximize existing members' welfare could actually increase the steady-state level of unionization. The model raises the possibility that legislation such as the Landrum-Griffin Act, which reduced incumbency advantages for union leaders, could have played a role in causing union decline. The dynamic nature of the model also suggests that increases in unionization may be rapid, whereas declines will occur more slowly, consistent with the empirical patterns of spurts in unionization, followed by a gradual decline, noted by Richard B. Freeman (1998).

By considering how union institutions evolve, this paper provides a new take on the old question of what unions maximize (e.g., John T. Dunlop 1944). The evolutionary perspective allows us to endogenize what unions maximize, and provides an explanation as to why what the unions we observe maximize is different from what one might expect based on the preferences of their members. Masahiko Aoki (1980) and Donald L. Martin (1980) argue that the rank and file may be more aggressive in their wage demands than would be optimal for the long-term survival of the union. This paper complements this earlier literature by showing how, in equilibrium, the union may evolve political institutions that allow the demands of the rank and file to be moderated and let the union implement the policies that are better for the union's long-term survival.

This paper also builds on earlier work that studies union organizing. William T. Dickens and Jonathan S. Leonard (1985), for example, show that unions must continually organize new enterprises in order to offset the natural decline in membership due to turnover among firms. Michael T. Hannan and John 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. This paper is also related to several papers that apply biological techniques to study firms, such as Richard R. Nelson and Sidney G. Winter (1982) and Prajit K. Dutta and Roy Radner (1999).¹

This paper can also be seen as fitting within the recent literature on institutions. A number of recent papers have documented the impact of institutions on economic performance, yet fewer papers address the determinants of institutional evolution and the question of why societies adopt institutions yielding suboptimal outcomes. We propose a model of institutional evolution under which unions with strong incumbency advantages spread at the expense of unions that better serve their members' interests.

¹ In Nelson and Winter (1982), evolutionary and maximizing models yield different dynamics, but the steady state predictions are similar, unlike in our model. Dutta and Radner (1999) argue 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 (1999) by explicitly modeling ongoing competition between unions and by modeling the spread of a union within a population of potential hosts. We therefore can explicitly derive evolutionarily stable behavior. We also demonstrate a mechanism, provisions for incumbency advantages in union constitutions, that has the equivalent function to an organism's genes, and test empirical predictions regarding union constitutions.

The remainder of the paper is organized as follows. Section I provides background on relevant US collective bargaining institutions. Section II presents the model and solves for the steady-state level of unionization with a single union and exogenous union wage and organizing policies. Section III shows that policies that maximize the welfare of workers will not be evolutionarily stable, that the evolutionarily stable union will either lower rent extraction or increase organizing expenditures, and that incumbency advantages for union leaders will be present in the evolutionarily stable union. Section IV presents empirical evidence. Section V concludes.

I. Background on US Collective Bargaining Institutions

Before introducing the model, it is useful to review a few features of US collective bargaining institutions relevant to the model. Outside of construction, music, and a few other industries, most new firms begin life without unions. Under the federal law covering most industries, if 30 percent of workers sign a petition calling for an election, the National Labor Relations Board (NLRB) conducts a certification election. It recognizes the union if more than half the workers vote for it.

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 by dues of existing union members, also play an important role.² 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, 1 in 20 workers who vote for a union in an organizing election are later found to have a valid claim for unfair dismissal by the NLRB (Paul 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 unions rarely switch affiliations, given the organizing costs involved and the AFL-CIO's constitutional prohibitions on member unions attempting to organize firms currently organized by different AFL-CIO member unions. The loss of existing union members is therefore

² For example, Paula Voos (1983) estimates that the marginal organizing expenditures required to win bargaining rights for an additional worker over the period from 1964 to 1977 were at least \$375 per person in 1980 dollars.

not primarily because of decertification elections, but because the firms covered by the union reduce employment or close down a unionized location.

The model in this paper is designed to apply to those US industries covered by the standard NLRB rules: new firms start as nonunion; 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. These simple institutional features create dynamics similar to those noted by Freeman (1998)—rapid increases in unionization in response to an increase in the ability of unions to organize or an increase in workers' desire to join unions—but much slower declines that occur through firm exit.

The dynamics of unionization levels also bear a similarity to those under the Susceptible-Infected (SI) model of epidemiological dynamics (see Roy M. Anderson and Robert M. 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. Note that this comparison is purely positive, not normative.

II. A Single Union with Exogenous Behavior

This section describes the basic model for the spread of a single union with exogenously given rent extraction and organizing expenditure. (Union behavior variables are endogenized in the next section.) Section IIA begins by outlining the entry, investment, and exit behavior of firms, taking union behavior as given. Section IIB describes how the union spreads and characterizes the steady-state level of unionization.

A. Firms

To generate the possibility that unions can extract rents, we assume firms have market power. In particular, we assume that firms produce one of a continuum of measure F possible products, and that there is a downward sloping demand curve, $q(p)$, for each product. Entry into a product market requires start-up costs. But once these costs have been paid, output is linear in labor and requires no other inputs, i.e. $q(L) = \psi L$. 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 firms is equal to F . For simplicity, we will assume that all firms face identical production functions, and so behave identically.

To pin down the wage in the absence of unions, and to ensure that union members value union jobs rather than simply assuming they will get another union job if their firm closes, we assume there is a competitive, constant return 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 that some are always employed in the home production sector, i.e., $W > L^*F$, where W is the quantity of workers, and L^* is the optimum quantity of workers each firm employs at wage \underline{w} .

In the absence of a union, each firm charges the profit-maximizing monopoly price P^* , pays workers the wage \underline{w} , and earns profits π , where profits are defined as

the surplus of revenues over the wages paid, $P^*q(P^*) - q(P^*)\underline{w}/\psi$. We assume that there is some demand for each product at a price above \underline{w}/ψ , so that each firm produces a positive amount, and that profits are maximized at some finite price.³

We assume that if the firm is unionized, the union extracts $\alpha + B$ from the firm, where α is additional wages for its members and B is available to the union to spend on organizing.⁴ Later, we will endogenize α and B , but from the perspective of the firm, these are exogenous parameters. For modeling convenience, we assume that 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.⁵ We assume that α and B are fixed amounts, independent of the size of the firms, so the unions do not distort the firm's decision about how much labor to hire. Thus, we model the impact of unions on employment solely on the extensive margin of increasing firm exit rates, not the intensive margin of reducing employment if the firm continues to exist. We conjecture that our results would generalize if there was an impact on the intensive margin as well.

Note that, in this model, π is exogenous to union density because there is only a single firm in each product market, and therefore each firm is already charging the monopoly price. If multiple firms compete in the same product market, then unions can increase π by reducing output and extracting the resulting product market rents (as, for example, in Michael Wallerstein 1989). In this model, then, organizing expenditures B should be thought of primarily as union organizing across markets, rather than within a given market, as we abstract from the question of multiple unions within a given market.

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 or investing in research and development to avoid being displaced by a competitor with superior technology.⁶ We assume that $\delta_I < 0$ and $\delta_{II} > 0$. When a firm exits another enters.⁷

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

⁴ Note that by assuming that the union can commit only to a constant level of rent extraction, we rule out two extreme cases. On the one hand, if the union had full powers of commitment, 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 firm's investments in staying alive. On the other hand, if the union had no ability to commit, the union would extract the full amount (i.e., set $\alpha = 1$) in each instant, and anticipating this, the firm would not invest at all. We intend this intermediate level of commitment on the part of unions to roughly correspond to the actual practice of negotiating medium-term contracts that periodically come up for renegotiation.

⁵ Some of our results might differ if unions have less bargaining power. In particular, if unions had less bargaining power, union members might elect leaders who have stronger preferences for higher wages than themselves so as to undo the effect of firm bargaining power and obtain the members' preferred outcomes. If this does not happen, and union leaders and firms bargain sincerely, then our model's predictions on wage bargaining become ambiguous, although unions with incumbency advantages will still potentially be able to outcompete those without such advantages in terms of attracting members. Michael Kremer and Benjamin A. Olken (2002) discuss in more depth a variant of the model in which firms have some bargaining power.

⁶ 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.

⁷ To close the model with a zero profit condition, one could assume that when a firm exits, an auction or, equivalently, a lottery, is held to determine what firm takes over production of the good. 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 a market

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

$$(1) \quad I(\alpha) = \arg \max_I \frac{\pi - \alpha - B - I}{r + \delta(I)}.$$

Investment is decreasing in rents extracted from firms, $\alpha + B$, since

$$(2) \quad \frac{dI}{d(\alpha + B)} = \frac{\delta_I}{\delta_{II}[\pi - \alpha - B - I]} < 0.$$

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

This argument about the relationship between unionization and investment is related to Paul A. Grout (1984), who shows that firms will invest less when there is a union that cannot make legally binding contracts. Robert A. Connolly, Barry T. Hirsch, and Mark Hirschey (1986) provide empirical evidence for this argument by showing that firms in industries where the unionization level was high tended to invest less in research and development. Hirsch (2004), in a survey of recent empirical work on unions, also concludes that unions may effectively tax the returns to long-lived capital investments, resulting in lower investment levels.

B. Steady-state Unionization Levels

Under the model, new firms are established without unions. Firms differ in how easily they are unionized, depending on factors ranging from the layout of the factory floor, to the personalities of managers, to the range of actions they can take to fight unionization. (In order to keep the model tractable, we consider a simple model in which firms, plants, and union bargaining units are coterminous.) We model this heterogeneity by assuming that each firm has a certain difficulty of being organized, which we denote by c (for cost). This cost c is a reduced-form way of capturing the many sources of firm heterogeneity in being unionized, including the degree to which firms will combat unionization. We will assume that for new firms c is distributed uniformly on the interval $[0, 1]$, and that c remains fixed for the life of the firm.

A union's organizing budget is equal to BU , where B represents the amount that unionized workers in each firm contribute toward the overall union's organizing

niche, as was widespread among Internet firms during the late 1990s. Assuming that there is competition among a large number of risk-neutral capitalists, the cost of entry will be equal to the expected value of owning a firm.

budget, and U is the number of unionized firms.⁸ In spending this budget, the union first targets those firms that are the easiest to organize (i.e., have the lowest c).

We assume that the attractiveness of unions to workers depends, among other things, on the amount of rents they extract for workers, α , and on the amount that the union spends on organizing other workers, B . Denote the attractiveness to workers of a union with a given level of α and B by $A(\alpha, B)$. Since workers recognize that firms will die off quickly if unions extract too much, $A(\alpha, B)$ generally will not be monotonic in α . Denote by (α_w, B_w) the level of rent extraction α and organizing expenditures B that maximizes a union's attractiveness to workers. For example, if workers cared only about the present discounted value of rents they receive, then $A(\alpha, B) = \alpha/[r + \delta(\alpha + B)]$ and $(\alpha_w, B_w) = (\arg \max \alpha/[r + \delta(\alpha)], 0)$. However, we allow this function $A(\alpha, B)$ to be much more general. In particular, we require only that A be continuous and differentiable, that $0 < \alpha_w < 1$, and that $B_w < 1$. For example, this would incorporate a case in which workers cared idealistically about contributing to the labor movement and had $B_w > 0$.

In addition to rent extraction α and organizing expenditures B , we assume that unions also differ on other characteristics that affect recruitment of new members, such as the type of worker they have experience organizing and the personality of the union leader. We assume that some unions are a better match for certain workers and other unions are a better match for other workers. This implies that $A(\alpha, B)$ is not discontinuous at (α_w, B_w) , but rather is continuous.

Given this measure of how attractive a union is to workers, $A(\alpha, B)$, we assume that the union's effective organizing budget is $A(\alpha, B)BU$. Thus, a union's organizing budget, BU , is augmented by how attractive the union is to potential members, $A(\alpha, B)$. The key assumption here is that there are two ways unions can attract workers. One is by making workers better off by choosing a combination of α and β that they prefer, i.e., by choosing a value of α and β that results in a higher $A(\alpha, B)$. The other is by spending money on organizing efforts that increase the likelihood workers will join unions through channels beyond simply their direct impact on the welfare of existing union members.

In steady state, then, in each instant a fraction $\delta(\alpha + B)$ of unionized firms exit, a fraction $\delta(0)$ of nonunion firms exit, new nonunion firms enter to replace these departing firms, and the union exhausts its effective organizing budget $A(\alpha, B)BU$, organizing some fraction of the newly entered firms.

Steady states of the system are characterized by p^* , the difficulty level below which all newborn firms are unionized, and U^* , the proportion of unionized firms. Since the distribution of unionization difficulties is uniform on $[0, 1]$, p^* is also the percentage of newborn firms that are unionized. In steady state, $U^* < p^*$, because unionized firms die at a faster rate than nonunion firms.⁹

⁸Note that this assumes that unions are limited in their ability to borrow to finance organizing efforts. We also assume that unions fully spend their organizing budgets each period rather than saving them to spend later. Fully spending the available budget would be optimal given standard preferences over the time path of unionization levels for unions that are either in steady state or approaching it from below. A union in steady state that experiences a negative shock that lowers its steady-state level of unionization might potentially want to delay spending its organizing budget until it faced an easier distribution of potentially unionizable firms.

⁹Of course, in the real world, factors outside the model may obscure this relationship. In particular, firms may differ in intrinsic profitability, and more profitable firms are more likely to attract attention from unions and less likely to exit.

The following proposition characterizes the steady state level of p and U :

PROPOSITION 1: *With a single union obtaining a level of rent α for workers and spending B on organizing, there can be two steady states: the trivial steady state with no unionization ($p^* = U^* = 0$), and the steady state with*

$$(3) \quad p^* = \begin{cases} \frac{2A(\alpha, B)B}{\delta(\alpha + B)} & \text{if } 2A(\alpha, B)B < \delta(\alpha + B); \\ 1 & \text{otherwise} \end{cases}$$

$$(4) \quad U^* = \begin{cases} \frac{2\delta(0) A(\alpha, B)B}{\delta(\alpha + B)^2 - 2A(\alpha, B)B[\delta(\alpha + B) - \delta(0)]} & \text{if } 2A(\alpha, B)B < \delta(\alpha + B). \\ 1 & \text{otherwise} \end{cases}$$

PROOF:

All proofs are in the Web Appendix available at <http://www.aeaweb.org/articles.php?doi=10.1257/app.1.2.150>.

Note that the nontrivial equilibrium is globally stable, and the trivial equilibrium with ($p^* = U^* = 0$) is unstable.¹⁰ From the expression for p^* , the “biological” intuition is apparent. In particular, p^* is increasing in the spread rate of unions and decreasing in the firms’ death rates. When $2A(\alpha, B)B \geq \delta(\alpha + B)$, the spread rate of the union to nonunion firms is substantial enough to overcome the death rate of unionized firms, so the stable steady-state unionization level is one. For the remainder of the paper we assume that

$$(5) \quad 2A(\alpha, B)B < \delta(\alpha + B),$$

for all α and B pairs such that $\alpha + B \leq 1$, unless otherwise stated, so that we are in the more interesting interior case with only partial unionization in the nontrivial steady state.

The following proposition states several comparative statistics about the equilibrium:

PROPOSITION 2: *Increases in $A(\alpha, B)$ increase steady-state unionization levels U^* . Increasing the death rate of all union and nonunion firms by a constant or by the same proportion reduces the steady-state level of unionization U^* .*

The intuition for the result about the relationship between death rates and unionization levels 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

¹⁰ The transition dynamics of the system, including the proof of global stability, can be found in the working paper version of the paper (Kremer and Olken 2002).

and less to expanding the size of the union. Empirical results supporting this implication of the model are presented in Section IV.

III. The Evolution of Unions

In biology, organisms with different genetic characteristics compete to occupy the same biological niches, and the organisms that survive are those that fare best in this competitive environment. In this section, we apply the same principles to the evolution of unions. First, we introduce into the model multiple unions that obtain different levels of rents for workers and spend different amounts on organizing. We characterize the wage and organizing policies that will be most successful in evolutionary competition, and argue that union leaders will only be able to implement these policies if they are protected by incumbency advantages embedded in union constitutions.

A. Competition Between Unions

When there are multiple unions, each would like to spend its organizing budget trying to organize the easiest firms first. Rather than assume that unions waste resources on battles to organize the same unorganized firms, we will assume that they divide firms so that at every level of difficulty, c , unions organize firms in proportion to their effective organizing budgets.¹¹ Since the effective organizing budget is the actual organizing budget (BU) multiplied by how attractive the union is to workers (indexed by the function $A(\alpha, B)$), unions that are more attractive to workers can organize disproportionately more firms. For example, suppose that there are two unions, 1 and 2, with union j having U_j member firms and implementing policies (α_j, B_j) . In that case, for firms that both unions wish to organize, they are divided up so that union 1 targets a fraction $A(\alpha_1, B_1)B_1U_1 / [A(\alpha_1, B_1)B_1U_1 + A(\alpha_2, B_2)B_2U_2]$ of them and the other union targets the remainder.

We can now identify the evolutionarily stable level of wages and organizing expenditures and show that it will differ from the welfare-maximizing levels.

PROPOSITION 3: *An entrant union can successfully invade and replace an incumbent union if and only if it has a higher value of $p^* = 2A(\alpha, B)B / \delta(\alpha + B)$.*

The key idea of the proof is that a union that can bear a higher average organizing 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

¹¹ We allow unions that extract more for their members to be more successful in attracting members, but rather than have 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. We conjecture that the main results of the paper (in particular, Propositions 4 and 5), would hold for any division of firms that is continuously increasing in the effective organizing budget at each difficulty level.

containing the union with the highest possible steady-state average organizing cost level, i.e., the highest level of equilibrium p^* .

The following proposition then follows immediately from Proposition 3.

PROPOSITION 4: *The union obtaining the level of rents α_S for its workers and having organizing expenditures B_S , where (α_S, B_S) maximizes p^* , will be evolutionarily stable. The union that chooses workers' preferred policies (α_W, B_W) will not be evolutionarily stable.*

Figure 1 gives a graphical explanation as to why $(\alpha_S, B_S) \neq (\alpha_W, B_W)$. The figure shows $A(\alpha)$, the attractiveness of the union to new members; $\delta(\alpha)$, the death rate of unionized firms; and $2A(\alpha)B / \delta(\alpha)$, the steady-state level of p^* , as functions of α , all drawn for $B = B_W$. The function $\delta(\alpha)$ 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. If one starts at the level of rent provision 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.¹² Therefore, the evolutionarily stable set of policies (α_S, B_S) must differ from workers' preferred policies.

This result holds as long as the attractiveness of unions to workers is continuous in α and B . If, for example, there were Bertrand competition among unions for potential members at unorganized firms, and workers always joined whichever union delivered greater discounted rents, then the slope of $A(\alpha)$ would be infinite at α_W , and the evolutionarily stable level of rent provision would equal the optimal amount of rent provision for current workers. If workers decide which union to join based not only on α and B but also on other idiosyncratic factors, however, union recruitment will vary continuously rather than discretely in α and B , and (α_S, B_S) will differ from (α_W, B_W) .

The model implies not only that (α_S, B_S) will differ from (α_W, B_W) , but also has implications for how the evolutionarily stable union will differ from the welfare maximizing union. The following proposition characterizes the relationship between (α_S, B_S) and (α_W, B_W) .

PROPOSITION 5: *The evolutionarily stable union either exacts less rent from firms than the welfare maximizing union ($\alpha_S + B_S < \alpha_W + B_W$), has higher union organizing expenditures ($B_S > B_W$), or both.*

To see the intuition, it is useful to think back to the analogy from epidemiology. Evolutionary fitness is determined by the spread rate of the organism on the one hand and the death rate on the other hand. Therefore, the evolutionarily stable union must be superior on one of the dimensions. Formally, note that Proposition 4 implies

¹² Note also that while reducing α in the neighborhood of (α_W, B_W) increases p^* and thus evolutionary fitness, there is no guarantee that the full (α_S, B_S) , which globally maximizes p , will have $\alpha_S < \alpha_W$. Proposition 5 characterizes the overall relationship between the policy pair (α_S, B_S) and (α_W, B_W) .

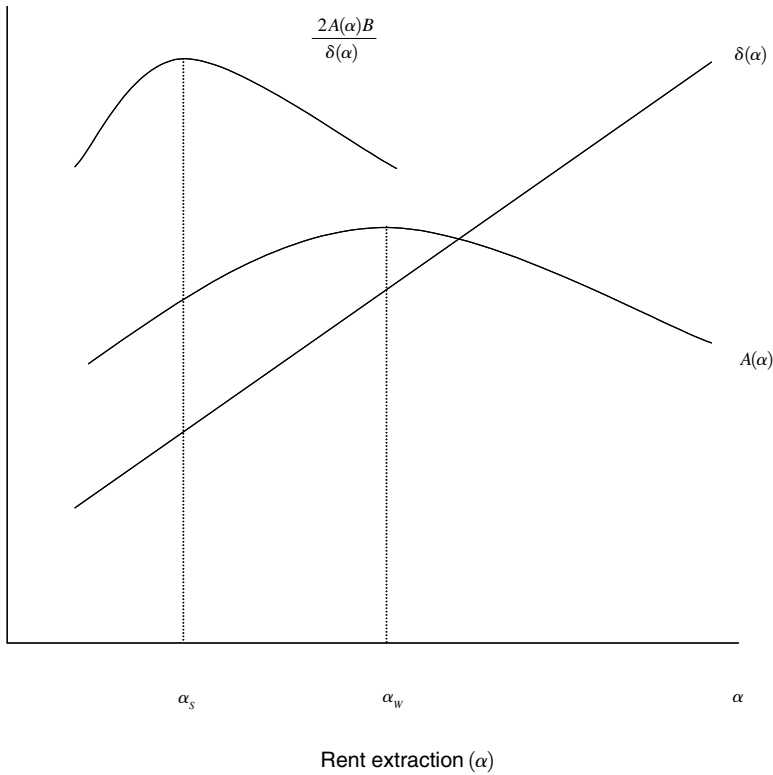


FIGURE 1. EVOLUTIONARILY STABLE AND WELFARE-MAXIMIZING LEVELS OF RENTS OBTAINED FOR WORKERS HOLDING B FIXED AT B_W

$p^*(\alpha_S, B_S)$ is strictly greater than $p^*(\alpha_W, B_W)$. From equation (3), it is clear that this can only be the case if death rates of unionized firms are lower or union spread rates are higher with (α_S, B_S) than with (α_W, B_W) .

Note that although (α_S, B_S) is evolutionarily stable and maximizes p^* , a single monopoly union could attain a higher level of U^* . The analogy to biology provides some intuition for this. Consider a biological mutation that allowed an organism to survive beyond reproductive age. This would increase the total number of organisms with the mutation that were alive at any moment (the equivalent of our U^*), but would not increase reproductive success, and hence would not convey an advantage in evolutionary competition. Trading off some lifespan for an increase in reproductive success would be evolutionarily advantageous.

In our case, policies that lengthen the life span of the firm, i.e., reducing α or B , will increase the life span of the union local in the firm as well. This has a direct effect on unionization levels. By extending the lives of unionized firms, it also has an indirect effect by generating more revenue to finance organizing drives. Only this second, indirect effect contributes toward evolutionary fitness. At (α_S, B_S) , small reductions in α or B will have only a second-order impact on fitness, but will have a first-order effect on the expected lifespan of existing unionized firms. This implies

that the level of rent extraction that maximizes U^* in the absence of competition from other unions will be less than the level of rent extraction that maximizes p^* .

The predictions of the model are less stark and more realistic when the model is extended to allow for the fact that unions sometimes cease to be able to organize workers, for internal reasons, such as the replacement of dynamic leaders by less competent ones, and due to changes in the external environment. In this more realistic setting, the system still selects for unions that create higher levels of p^* than the welfare-maximizing union. The system does not fully converge to the α_S union as it did above. Instead, the steady state is a distribution of unions that is skewed toward values of (α, B) that have higher p^* than the values of (α, B) in the distribution of newly born unions. (See Kremer and Olken 2002.)

B. Implications for Union Institutions

Proposition 4 showed that the evolutionarily stable union must follow policies that would not be chosen by its members democratically. If all unions were controlled by their rank and file, they would choose rent provision policies that maximize worker welfare. For the model to be applicable, there must be a mechanism that allows some unions to systematically deviate from the policies preferred by members. Moreover, this mechanism must persist over time.

Provisions of union constitutions, in particular provisions creating incumbency advantages for union leaders, can provide a persistent mechanism influencing union behavior. While animals pass on their genes to offspring, thus influencing their behavior, union members at one firm pay dues which finance organizing drives at other firms. When organizing drives are successful, workers at the newly organized firm are governed by the same national union constitution as the union members who financed the organizing drive, even if the newly organized members are in a separate local. The union constitution and the national union institutions are the transmission mechanism through which union characteristics persist over time. Moreover, while union constitutions can, in theory, be amended, in practice they are highly persistent over time, like genes.¹³

Union constitutions can contain a variety of provisions that provide advantages to incumbents seeking to run for reelection. In particular, many unions have constitutional rules specifying indirect leadership elections in which the president is elected by delegates to a national convention rather than by the membership at large. These indirect elections are often combined with provisions allowing national leaders to put locals in trusteeship and giving them control over strike funds. Since convention delegates are typically local union leaders, they face strong pressure to support incumbents in national office if they think that the incumbents will win (Thomas

¹³ To examine this, we randomly selected 20 unions from our dataset and compared the constitutional provisions for electing the union president that were in force in 1955 with those in force in the most recent available constitution. For 19 of the 20 unions, indirect elections—the key provision of incumbency advantages we examine in the empirical work below—stayed constant over this nearly 50 year period. The only 1 of the 20 unions for which they had changed was the Teamsters, where direct elections for union president were imposed as part of the federal government takeover of the union.

Geoghegan 1992 and Herman Benson 1986).¹⁴ Coordinating to overcome the free-rider problem in risking the private cost of retaliation for the public good of better leadership is difficult, particularly in large unions.

Incumbency advantages create the potential for persistence in union policies that depart from the preferences of the rank and file, particularly if union leaders have influence over their successors and tend to choose successors with similar preferences.¹⁵ Suppose that unions obtain their first leader from a distribution of potential leaders with different preferred policies, and that the support for these preferred policies includes (α_S, B_S) . For example, different potential union leaders may have different ideological views. Suppose also that in the absence of incumbency advantages, leaders are constrained to implement the policies preferred by workers, but that with incumbency advantages, they can pursue their own preferred policies. In this simple context, some of the unions with strong incumbency advantages will do quite badly, but others will do well. As time goes on the unions that do badly will disappear from the scene while the ones that are close to the evolutionarily stable level of α and B will survive and grow at the expense of unions without incumbency advantages and unions with incumbency advantages and evolutionarily disadvantageous choices of α and B .

Of course, potential leaders may be particularly likely to have certain specific preferences over α and B . In particular, many have argued that union leaders are motivated to build up their organization (Robert Michels 1949; Herbert A. Simon 1947). If some union leaders seek to maximize the long run size of their union and are protected by incumbency advantages, their unions will survive, and others will be displaced. Note, however, that our model does not rely on the assumption that there is a universal tendency of leaders to have these preferences. All that is required is that some proportion of leaders act this way, and that there is a transmission mechanism within unions such that successors are inclined to follow policies similar to those of their predecessors.

Finally, one might imagine that some union leaders are simply opportunists who offer to reduce rent extraction from firms in exchange for private benefits from firm owners. For example, union leaders might offer to extract less from firms in exchange for private benefits for themselves. These private benefits can be monetary offers in the form of corruption or, as more frequently happens, contract provisions that are of particular benefit to union leaders. For example, Arthur M. Ross (1950) argues that unions often make concessions on rank-and-file 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. In Kremer and Olken (2002), we consider a simple case in which organizing expenditures are either exogenous or zero for the welfare-maximizing union, and we show that, in this case, unions led by opportunistic

¹⁴ Incumbency advantages arise from other union rules as well. For example, union rules often do not prohibit union staff from supporting reelection campaigns, put local union officers in charge of vote counting in union elections, and allow incumbent leaders to promote their candidacies in union newsletters and to deny challengers access to membership lists or even contact information for local chapters.

¹⁵ This could happen in many ways. For example, some unions have informal family dynasties in which union leadership has been passed down from father to son.

leaders with incumbency advantages will have an evolutionary advantage relative to unions that maximize the welfare of members. To see that an optimistic union leader with some incumbency advantages will pursue more evolutionarily fit policies than one without evolutionary advantages, note that if opportunists collude with firms to reduce α slightly from (α_w, B_w) , there will be a second-order impact on union spread rates but a first-order increase in firm lifespan.

IV. Evidence

The model implies that a population of unions following workers' preferred wage and organizing policies can be successfully invaded by unions extracting fewer rents from firms or spending more on organizing. As a result, incumbency advantages allowing union leaders to depart from members' preferred policies should become widespread, and unions without such incumbency advantages should be more likely to decline. Proposition 2 also implies that since overall union prevalence depends on the death rate of firms δ , union concentration should be higher in industries with lower death rates of firms. This section examines empirical evidence for these predictions.

A. Incumbency Advantages and Union Decline

We test the key empirical prediction of the model—that unions with constitutional incumbency advantages for union leaders will outcompete unions without such provisions—using data on American union membership since the AFL-CIO merger of 1955.¹⁶ Since the model is designed for private sector firms, we limit the sample to private-sector unions, though the results are similar if public-sector unions are included as well. For each private-sector union that existed at the time of the AFL-CIO merger, we obtained the national union constitution that was in force at that time.¹⁷ As discussed in Section IIIB, one of the most important sources of incumbency advantages is indirect leadership elections. We therefore code each union based on whether the constitution allows for direct or indirect elections of the union president. In 1955, about 20 percent of unions had direct elections for union president.

We also obtain membership data from 1955–1995 for each union. Over this time, many small, declining unions were absorbed into larger unions. When two unions merge, we attribute the membership of the newly created union back to the original unions that formed it, according to their relative sizes at the time of the merger.

For each union, we computed the difference between the log of the 1995 membership and the log of the 1955 membership. Figure 2 plots a kernel density of this measure of union growth separately for unions with direct elections and unions with

¹⁶ 1955 is the earliest starting date for which we could obtain consistent union membership data.

¹⁷ The sample includes all unions that are, or ever were, members of the AFL-CIO for which we could obtain constitutions, a total of 130 unions. In 20 percent of cases (27 unions), we were unable to obtain the constitution in effect in 1955, and therefore used the earliest constitution we could obtain, the vast majority of which were from before 1960. When we drop unions for which we could not obtain the pre-1955 constitution from the sample, results are similar.

indirect elections. As is apparent from the figure, the distribution of union growth is substantially shifted to the left for unions with direct elections, which implies that unions with direct elections were substantially more likely to decline than unions with indirect elections.¹⁸ In fact, in 1955, 29 percent of members were in unions with direct elections for union president, whereas, in 1995, this figure had fallen to 19 percent.

To examine this relationship in a regression framework, we code a union as having experienced a substantial decline if its membership declined by 50 percent or more from 1955 to 1995, or if it ceased to exist as a result of a merger.¹⁹ This captures the fact that unions decline either by shrinking in membership, or by being absorbed into another union, in which case their constitutional provisions are lost.²⁰ We then estimate the following equation by OLS:²¹

$$(6) \quad \text{DECLINE}_{ij} = \alpha_1 + \alpha_2 \text{DIRECT}_{ij} + \alpha_3 \text{STRIKEVOTE}_{ij} + \alpha_4 X_{ij} + \gamma_i + \varepsilon,$$

where i represents the primary one-digit Standard Industrial Classification (SIC) industry where the union is active, j represents a union, *DIRECT* is a dummy for the union having direct elections for union president, X are other union characteristics (such as the size of the union in 1955 and whether the union was created as part of the CIO), and γ_i are a set of industry fixed effects, to capture the fact that the shocks experienced by different industries may have been different. At the local union level, one provision that could restrain radical behavior is the difficulty of calling strikes.²² For each union, we therefore coded whether a strike could be authorized with a

¹⁸ Interestingly, Figure 2 also suggests that the variance of union growth is smaller in unions with direct elections than in unions with indirect elections. As noted above, indirect elections allow for both leaders who maximize union growth as well as destructive leaders who lead to the union's decline, whereas direct elections reduce both extremes. This is consistent with the findings of Benjamin F. Jones and Olken (2005), who show that, among country leaders (i.e., presidents and prime ministers), there appears to be higher variance in leader quality among autocrats than among democrats.

¹⁹ Results where we set the threshold for decline at 75 percent or 25 percent, instead of the 50-percent threshold used for the results in the table, are qualitatively similar to the results presented. We followed the Bureau of National Affairs (2001) in their classification of which unions ceased to exist as a result of a merger. Generally, if a small union merged with a much larger union, then only the small union would be coded as ceasing to exist; if two unions of approximately equal size merged to create a new union, both unions would be coded as ceasing to exist.

²⁰ It is important to include these two different ways in which unions decline, because using the imputed growth measure alone will not necessarily capture decline. For example, consider a small union that was declining in the 1950s and 1960s, but then merged into the larger Service Employees International Union (SEIU), which experienced high growth over the subsequent period. If we only examine the union's imputed growth, we may find an increase, even though the union declined and its constitutional provisions were lost. In results available from the authors on request, we have separately estimated equations for whether a union exited from the sample between 1955 and 1995 and change in log union size between 1955 and 1995 conditional on not exiting. We find that the negative effect of direct elections operates through exiting the sample rather than changes in union size conditional on remaining in the sample.

²¹ Results are virtually identical to marginal coefficients calculated from estimating columns 1–3 (i.e., nonfixed effects specifications) using a probit specification instead.

²² While constitutional rules requiring supermajorities may tend to moderate wage demands, these rules could also restrict the total resources coming into the union and also could help reduce organizing expenditures, particularly if workers' willingness to pay dues is tied to the amount the union delivers to them on wages. The model, therefore, does not deliver unambiguous predictions about the impact of strike-vote provisions in union constitutions on the spread of union growth.

TABLE 1—DETERMINANTS OF UNION DECLINE

	(1)	(2)	(3)	(4)
Direct	0.087 (0.094)	0.173* (0.091)	0.184* (0.095)	0.199** (0.094)
Supermajority strike vote rule	0.085 (0.086)	0.061 (0.083)	0.072 (0.088)	-0.107 (0.089)
1995 Union membership (thousands)		-0.001*** (0.000)	-0.001*** (0.000)	-0.000** (0.000)
CIO			0.001 (0.102)	-0.152 (0.099)
Constant	0.653*** (0.060)	0.732*** (0.061)	0.719*** (0.068)	0.806*** (0.061)
1-Digit SIC code fixed effects	N	N	N	Y
Observations	116	110	106	106
R ²	0.02	0.10	0.09	0.28

Notes: Dependent variable is a dummy for whether union membership declined by 50 percent or more from 1955–1995 or whether the union ceased to exist as a result of a merger over the same period. Linear probability model. Robust standard errors in parentheses.

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

majority vote of the local union, or whether such a vote requires a supermajority (*STRIKEVOTE*).²³

The results are presented in Table 1. The results confirm statistically what is shown in Figure 2—incumbency advantages substantially decrease the probability of union decline. Specifically, column 4, which includes one-digit industry fixed effects, indicates that unions with direct elections are 20 percentage points more likely to have declined than unions with indirect elections. Note that the results are substantially stronger in columns 2–4, where we include a control for the initial size of the union in 1955. If there are any fixed costs to running a union, then unions that are larger initially are less likely to exit, as they can still afford to shrink substantially without exiting. Therefore, controlling for initial size separates out this effect from direct elections, and strengthens the results. In results not shown in Table 1, we also find that the effect of direct elections on union decline is stronger among unions that were larger in 1955, consistent with the idea that indirect elections provide more of an incumbency advantage in larger unions.

Union constitutions are very persistent, so it is not possible to use panel-data techniques to repeat the above analysis focusing only on changes in union constitutions. However, over the period we examine, there are a few cases of substantial, plausibly exogenous increases in democracy. In the 1970s, the head of the United Mine Workers (UMW) arranged for the murder of the opposition candidate and his family. In response, the federal government stepped in to oversee the next election. Dissident movements within the UMW grew, wage demands escalated, and

²³ For the unions where this provision was not explicitly specified, we assume that the majority vote is sufficient to call a strike. Of unions that require a supermajority, virtually all require either 2/3 or 7/10 vote.

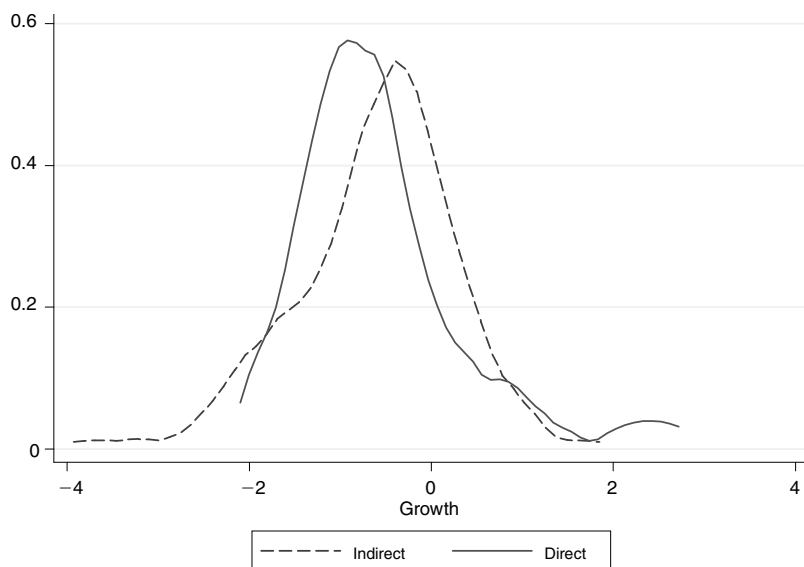


FIGURE 2. UNION GROWTH 1955-1995

UMW membership collapsed.²⁴ More recently, the federal government takeover of the Teamsters was followed by contested elections in which both sides promised greater militancy on wages. Teamster membership, which had been growing previously, began to decline.²⁵ Though these are only two cases—not enough for a regression-based analysis—these declines associated with increased democracy are consistent with the cross-sectional results presented in Table 1.

Note that a simple model in which all leaders are empire builders might suggest that unions with incumbency advantages would grow faster than all unions without such advantages. By contrast, an evolutionary model suggests that incumbency advantages would lead some unions with incumbency advantages to grow rapidly (in particular, those unions with leaders whose preferences are close to (α_S, B_S)), but would also allow some unions with incumbency advantages to decline rapidly (in particular, those unions with leaders whose preferences were very far from (α_S, B_S)). By contrast, democratic unions should all follow policies (α_W, B_W) , and their membership should decline more uniformly. In fact, the evidence is consistent with this. During the period we study empirically, the standard deviation of log union size grew from 1.69 to 1.94 among unions with indirect elections and only 1.62 to 1.75 among unions with direct elections.

Given that this evolutionary process has been underway for many years, one would expect most of today's unions to exhibit substantial incumbency advantages. Suggestive evidence for this is provided by the low rate of election defeats for incumbent union

²⁴ This also sparked a strong dissident movement in the Steelworkers, who had strong historical ties to the UMW. The dissident movement may have played a role in the similar escalation of wage demands in the steel industry, and potentially in the collapse of steel employment.

²⁵ The federal government has more recently intervened at the national level for the Laborers and the Longshoremens, but as the Laborers' longtime leader was removed in 2000, and the government only took over the Longshoremens recently, it is too early to assess the long-term impact of these changes.

leaders in today's large unions. Of the ten largest US unions, we could obtain complete data for nine. Over the history of these unions, an incumbent union president had only a 0.6 percent chance of being defeated in an election each year. By way of comparison, during the twentieth century, an incumbent US president had a 5 percent annualized chance of being defeated in a general election, and during the period from 1950–2000 members of the House of Representatives had a 2.5 percent annualized chance of being defeated in an election.²⁶ This suggests that the incumbency advantages enjoyed by today's union presidents are substantial, consistent with decades of evolutionary competition favoring unions with greater incumbency advantages.

B. Firm Turnover and Unionization Levels

Proposition 2 of the model implies that unions should be more prevalent in industries with exogenously lower firm turnover. For example, hotels should have higher unionization rates than restaurants. While this prediction is not unique to this model (for example, heterogeneity in rents across industries could also generate a correlation between firm exit rates and unionization rates), it is a testable implication of the model, and seems worth examining in the data.

To test this prediction, we use Census of Manufactures data from Timothy Dunne, Mark J. Roberts, and Larry Samuelson (1988) on firm exit rates across four-digit US manufacturing industries. We focus on their measure of exit rates for firms producing 99 percent of industry output (*EXIT*), though results are similar if we use the measure that includes all firms.²⁷ We use union membership rates, defined as the percentage of employees in an industry who are union members, compiled by Hirsch and David A. MacPherson (1993) from the Current Population Survey for 231 three-digit Census of Population industries (*UNIONMEMBERSHIP*). The precision of our measures of union density varies across industries, since the number of observations from the CPS for each industry varies from 5 to 21,950. Therefore, we weight each industry by the square root of the number of observations.

A number of other theories might suggest a correlation between unionization and other variables which might be correlated with firm exit rates, so we control for the Herfindahl-Hirschman index of concentration (*HHI*), average firm size (*SIZE*), and capital intensity (*K*) using data from the 1992 Census of Manufactures on four-digit SIC manufacturing industries. After combining the different datasets, we have data for 66 Census of Population industries. We then estimate the following regression:

$$(7) \quad \text{UNIONMEMBERSHIP}_i = \alpha_1 + \alpha_2 \text{EXIT}_i + \alpha_3 \text{HHI}_i + \alpha_4 \text{SIZE}_i + \alpha_5 \text{K}_i + \varepsilon_i,$$

where i is an industry.

²⁶ The probability that the president was defeated in election was calculated by the authors. The probability a member of the House of Representatives was defeated was calculated by averaging values reported in Alan I. Abramowitz, Brad Alexander, and Matthew Gunning (2006).

²⁷ Note that one might be concerned about potential endogeneity in turnover rates if, for the reasons discussed above, unions cause higher firm death rates. However, this would cause a bias in the opposite direction of the results we find in Table 2. Given these results, it seems reasonable that differences in firm exit rates across industries might be primarily due to exogenous industry characteristics, though differences in exit rates within industries might be more closely associated with union behavior.

TABLE 2—UNIONIZATION AND FIRM TURNOVER

	(1)	(2)	(3)	(4)	(5)
Firm exit rates	-3.539*** (0.628)	-3.056*** (0.658)	-3.941*** (0.754)	-3.528*** (0.779)	-3.438*** (0.871)
Herfindahl index of industry concentration		60.652** (29.851)			92.438*** (33.007)
Average firm employment			-0.010 (0.010)		-0.022** (0.011)
Capital intensity (i.e., ratio of assets to employment)				0.001 (0.032)	0.019 (0.031)
Constant	52.655*** (4.954)	44.928*** (6.153)	56.705*** (6.502)	52.539*** (6.914)	47.151*** (8.550)
Observations	66	66	66	66	66
R ²	0.33	0.37	0.34	0.33	0.42

Notes: Dependent variable is percent of employees in each industry in a union (i.e., 0–100). Exit rate is the probability a firm in a given industry exists each year, also scaled 0–100. Each observation is weighted by the square root of the number of observations from which the union membership average was constructed. Standard errors are in parentheses.

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

The results are presented in Table 2. As the model suggests, higher firm turnover rates—i.e., greater values of δ —are associated with lower union membership rates.²⁸ In our preferred specification, shown in column 5 of Table 2, a 1 percentage point increase in the exit rate is associated with a 3.4 percentage point decrease in the unionization rate. Interestingly, one implication of the results is that the high unionization rates among public sector institutions and among large firms may be due, in part, to the low exit rates in these institutions. Private-sector unions have to constantly organize new firms just to keep membership constant, whereas once a public-sector union organizes a new working unit in the government, it can be assured of new members for some time to come (barring extreme events like the Professional Air Traffic Controllers (PATCO) strike).

C. Longer Term Evidence from American Union History

Although we can only provide quantitative evidence for the postwar period, the longer-run history of American unionism is also consistent with the model's predictions that over time the union movement should become more centrally controlled and less representative of the rank and file. Prior to the 1880s, most American unions were local organizations, and union leadership was close to the rank and file. In the early 1880s, many of these local unions were affiliated with the Knights of Labor, the first large-scale national labor organization (Kim Voss 1993). The locus of power in the Knights was the district assembly, whose residents were usually from the same city and

²⁸ In results not reported here, we also find similar results when we examine union coverage rather than union membership.

were relatively independent of the more moderate national union (Lloyd Ulman 1955). Unrestrained by the national leadership, Knights locals appear to have overreached. They lost a number of major strikes and sparked heavy opposition by industrial interests. Unlike labor relations in later periods, very few strikes in the period when the Knights were preeminent ended in compromises; instead, strikes tended to be all-or-nothing affairs that were either won or lost (David Card and Craig A. Olson 1995).

As the Knights began to lose major strikes called without the national leadership's approval, they rapidly declined and lost their place as the dominant labor organization in the United States to the AFL, which was founded in 1886. In contrast to the Knights, the AFL was founded, in part, on the principle of switching the locus of power from district assemblies to national trade unions. Since national unions were spread out across the country, members would have had substantially less grassroots input into decision making and ability to coordinate challenges to incumbents.²⁹

Founded in 1935, the CIO took a different, even more centralized approach to union organizing. Other than the original founding unions, the majority of CIO unions were formed from organizing committees set up and funded by unions from other industries to recruit workers in new industries. Cross-industry organizing of this type would be difficult to explain using many conventional models, but would be consistent with the evolutionary story presented here. Because CIO unions were organized from the top down rather than from the bottom up, and because they tended to bargain with large, national corporations, CIO unions were constructed to retain more central control over bargaining than AFL unions. For example, the national leadership of the Steelworkers and the UAW used their constitutional power to suspend officials of local unions to ensure consistent policies across locals (Seymour M. Lipset 1960), and the national president of the Steelworkers union was the same person the CIO had appointed to head the organizing committee, rather than a rank-and-file Steelworkers' union member.

The model would predict that CIO unions, with their top-down organization and emphasis on national control, should be able to successfully invade the AFL unions, where control was more local. In fact, the CIO did grow explosively. This was due, in large part, to the active efforts of the CIO to organize new members, an effort which consumed half the CIO budget during its first three years (Philip Taft 1964). By the time of the merger with the AFL in 1955, only 20 years after its founding, the CIO accounted for more than 5 million members. While there are other explanations that may account for the rise of the CIO, such as the switch to an industry-based model of unions rather than a craft-based model, the link between increased spending on organizing (increases in B) and centralization of control is consistent with the theory.

In 1955, competition between the AFL and CIO was brought to a close by the AFL-CIO merger. The new AFL-CIO anticompetitive agreement can be considered a decrease in the degree of union entry. Moreover, it also restricted the ability of CIO unions, which had higher organizing expenditures, to expand into industries that had

²⁹ Of course, other possible explanations have been offered for the decline of the Knights and the rise of the AFL, most notably the organizational difficulties in the Knights caused by the inclusion of nonworkers in so-called "mixed lodges" and the need for national coordination of policies provided by AFL trade unions (Ulman 1955).

an AFL union presence. In 1959, the passage of the Landrum-Griffin Act reduced incumbency advantages by regulating trusteeships and by mandating increased union transparency and regular elections. In accounts of US labor history, the Landrum-Griffin Act typically does not receive as much attention as other legislative changes, such as the Taft-Hartley Act, but our analysis suggests that it may have considerable hidden power. To the extent that this law reduced incumbency advantages below the evolutionarily stable level, Proposition 4 predicts that the law would have increased welfare for current unionized workers but also been associated with a decline in unionization. During this period, union density has been declining, from a high of 35 percent of the nonagricultural workforce in 1945 to 13 percent of the workforce in 2000. Of course, many other reasons have been proposed for the decline in union density in the United States since 1955, but the increase in union democracy caused by the Landrum-Griffin Act, and the reduction in the ability of CIO unions to organize in areas where AFL unions had jurisdiction, may have contributed at least to some extent.³⁰

During this period of overall union decline, one of the most rapidly spreading unions was the SEIU, which doubled in size over the past 20 years, from 688,000 members in 1985 to 1,354,000 members in 2004. Moreover, most of this growth came from its own organizing efforts, rather than from mergers, and a substantial fraction of that organizing effort was directed at workers whose firms did not compete directly with those of existing union members. For example, it is not clear how janitors in Boston benefit from organizing janitors in Houston. Consistent with the model, the SEIU is a highly centralized union, which our model suggests would be necessary to sustain such high levels of organizing expenditures *B*.

V. Conclusion

This paper has applied techniques from biology to model the evolution of union institutions. We show that unions that choose the optimal wage and organizing policies for their members will be displaced in evolutionary competition by unions that either extract less from firms, allowing them to live longer, or spend more on union organizing, or both. For union leaders to deviate from workers' preferences, they must be insulated by incumbency advantages. The model predicts that unions with these incumbency advantages should displace unions without them.

As discussed above, the model is consistent with a number of stylized facts: the unions we observe today exhibit substantial incumbency advantages; unions with fewer incumbency advantages have been more likely to decline; and when federal intervention decreases incumbency advantages, wage demands have increased and unions have declined.

In addition, the model can help explain two more stylized facts that pose puzzles for standard models. First, challenges to incumbents from rank-and-file dissidents

³⁰ Other explanations include sectoral shifts away from traditionally unionized sectors (Henry S. Farber 1985, Dickens and Leonard 1985, Leo Troy 1990, 1992), demographic shifts, particularly the increase in female and part-time employment (Robert J. Gordon 2002), improved provision of benefits by firms and government (George R. Neumann and Ellen R. Rissman 1984), increased management opposition to unions (Weiler 1983 and Freeman 1985, 1998), and changes in worker attitudes toward unions (Lipset 1986).

almost always press for higher rent extraction, rather than arguing that incumbents are too extreme. By contrast, under a simple median voter model, one might expect that half of union members would prefer a more aggressive approach to wage bargaining than leaders would implement, and half would prefer a more moderate approach. Second, 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 the United Auto Workers organize graduate students at New York University. It is unclear that steel workers or auto workers benefit sufficiently from organizing restaurant workers and college students to justify the expense involved.³¹

The model outlined in this paper is consistent with agency-theory models in which union leaders are agents whose interests differ from those of their principals, the rank and file. It is also consistent with sociological theories such as Michels’ (1949) “Iron Law of Oligarchy” which suggests that leaders generally seize control of their organizations and work to preserve the organization itself rather than advance the original goals of the organization. These considerations may well be the proximate cause of moderation of wage demands by union leaders. However, under standard agency models, principals design optimal mechanisms for agents. Agency theory thus begs the question of why so many existing 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, and, hence, we will observe many of these unions. Similarly, the model here suggests that 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.

The theory presented in this model differs from a traditional theory of empire-building, in that we do not assume that all union leaders are empire builders. Instead, our model suggests that there will be a wide distribution of union leaders, each with different preferences and proclivities. Over time, those union leaders that maximize evolutionary fitness—the ability to out-compete rival unions—will survive, rather than the union leaders that maximize the size of their union. Moreover, incumbency advantages will not lead uniformly to growth in union size. Rather, some unions with incumbency advantages will do well, and others will do poorly. Over time, the unions with incumbency advantages and union leaders who maximize evolutionary fitness will outcompete those unions either with fewer incumbency advantages or unions with incumbency advantages but union leaders who follow other sets of policies.

In the model, we adopt an explicitly dynamic approach in which the level of unionization is a state variable, and changes in unionization levels are driven by the relative balance of rates at which new firms are unionized and existing unionized firms exit. This framework matches the empirical asymmetry between patterns

³¹ More members may increase the political power of the union, but they may also demand a share of the fruits of that power.

of union growth and union decline, with union declines occurring gradually over time and union growth occurring in bursts (Freeman 1998), because the rate of firm death creates an upper bound on the rate of deunionization, while there is no such natural upper bound on the rate of increase in unionization. Moreover, in dynamic models such as ours, relatively small shifts in the spread rate of unions can have large long-run consequences, since the number of firms newly unionized each period depends positively on total union organizing expenditure and thus on the existing number of unionized workers. This class of dynamic models may therefore be useful in explaining the long-run decline in unionization rates in the United States.

The normative implications of the analysis are unclear, as it is ambiguous whether the welfare-maximizing or evolutionarily stable union will provide more total rents to workers, given the fact that unionization rates may be higher under the evolutionarily stable union. The impact on productivity depends on how we interpret unobservable investment by firms and start-up expenditures.

Regardless of these general equilibrium effects, however, the model implies that unions are not obtaining the optimal level of rent for their workers. Changing union constitutions to reduce incumbency advantages will likely increase welfare for the union's current members, but any reforms that decrease incumbency advantages from the evolutionarily stable level for a particular union will only be temporary, as unions that undertake such reforms will be displaced by unions that have more incumbency advantages. Changes to government rules that govern all labor organizations, such as the minimum union democracy standards required by the US Landrum-Griffin Act, are likely to be more lasting.

The analysis in the rest of the paper assumes that union constitutions are permanent and that successors take over the policies of their predecessors, so that the wage and organizing policies of unions are persistent. We presented evidence that constitutions, in fact, change only rarely, and there is certainly likely to be a fair amount of continuity in policies. However, changes do occur. One possible extension of the approach here would be to consider a model in which the strength of incumbency advantages changes occasionally. We conjecture that such a model would have continuous turnover in unions. Small unions with incumbency advantages and strong leaders would grow. Eventually, however, constitutions might be changed to weaken incumbency advantages or, alternatively, the dynasty of a particular type of leader might simply lose power to others. At that point, the large unions would begin to stagnate and decline, only to be replaced by aggressive newcomers.

Although this paper has focused on unions, similar evolutionary arguments could be made about other types of organizations. For example, 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 is optimal for the universities' educational and research mission. Those religions that grow may be those that are most successful at retaining members, rather than those that maximize members' welfare. 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.

Reality is likely to fall between the predictions of models in which institutions maximize their owners' welfare, and biological models in which organizational

characteristics are fixed. The more members have opportunities to control their organizations, the closer reality is likely to be to the welfare-maximizing model. “Mutation” is much more common in human institutions than biological organisms. However, human institutions also differ in the ease with which they can be restructured. Firms, which are subject to outside takeover and which are often controlled by a few large shareholders, may be reasonably well approximated by models in which the welfare of the principal is maximized. Unions, with diffuse control rights and little threat of external takeover, are less constrained to serve their members, suggesting that the difference between evolutionarily stable and welfare-maximizing behavior may be substantial.

REFERENCES

- Abramowitz, Alan I., Brad Alexander, and Matthew Gunning.** 2006. “Incumbency, Redistricting, and the Decline of Competition in U.S. House Elections.” *The Journal of Politics*, 68(1): 75–88.
- Anderson, Roy M., and Robert M. May.** 1991. *Infectious Diseases of Humans: Dynamics and Control*. New York: Oxford University Press.
- Aoki, Masahiko.** 1980. “A Model of the Firm as a Stockholder-Employee Cooperative Game.” *American Economic Review*, 70(4): 600–10.
- Benson, Herman.** 1986. “The Fight for Union Democracy.” In *Unions in Transition: Entering the Second Century*, ed. Seymour Martin Lipset, 323–73. San Francisco: Institute for Contemporary Studies.
- Bureau of National Affairs.** 2001. *Directory of U.S. Labor Organizations*. Washington, DC.
- Card, David, and Craig A. Olson.** 1995. “Bargaining Power, Strike Durations, and Wage Outcomes: An Analysis of Strikes in the 1880s.” *Journal of Labor Economics*, 13(1): 32–61.
- ▶ **Connolly, Robert A., Barry T. Hirsch, and Mark Hirschey.** 1986. “Union Rent Seeking, Intangible Capital, and Market Value of the Firm.” *Review of Economics and Statistics*, 68(4): 567–77.
- ▶ **Dickens, William T., and Jonathan S. Leonard.** 1985. “Accounting for the Decline in Union Membership, 1950–1980.” *Industrial and Labor Relations Review*, 38(3): 323–34.
- ▶ **Dunne, Timothy, Mark J. Roberts, and Larry Samuelson.** 1988. “Patterns of Firm Entry and Exit in U.S. Manufacturing Industries.” *RAND Journal of Economics*, 19(4): 495–515.
- Dutta, Prajit K., and Roy Radner.** 1999. “Profit Maximization and the Market Selection Hypothesis.” *Review of Economic Studies*, 66(4): 769–98.
- Dunlop, John T.** 1944. *Wage Determination Under Trade Unions*. New York: Macmillan Company.
- Farber, Henry S.** 1985. “The Extent of Unionism in the United States.” In *Challenges and Choices Facing American Labor*, ed. Thomas A. Kochan, 15–43. Cambridge, MA: MIT Press.
- Freeman, Richard B.** 1985. “Why Are Unions Faring Poorly in NLRB Representation Elections?” In *Challenges and Choices Facing American Labor*, ed. Thomas Kochan, 45–64. Cambridge, MA: MIT Press.
- Freeman, Richard B.** 1998. “Spurts In Union Growth: Defining Moments And Social Processes.” In *The Defining Moment: The Great Depression and the American Economy in the Twentieth Century*, ed. Michael D. Bordo, Claudia Goldin and Eugene N. White, 265–96. Chicago: University of Chicago Press.
- 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.
- Gordon, Robert J.** 2002. Discussion of Daron Acemoglu, Philippe Aghion, and Giovanni L. Violante, “Deunionization, Technical Change, and Inequality.” <http://faculty-web.at.northwestern.edu/economics/gordon/Crptpub.pdf>
- ▶ **Grout, Paul A.** 1984. “Investment and Wages in the Absence of Binding Contracts: A Nash Bargaining Approach.” *Econometrica*, 52(2): 449–60.
- Hannan, Michael T., and John Freeman.** 1987. “The Ecology of Organizational Founding: American Labor Unions, 1836–1985.” *American Journal of Sociology*, 92(4): 910–43.
- Hannan, Michael T., and John Freeman.** 1988. “The Ecology of Organizational Mortality: American Labor Unions 1836–1985.” *American Journal of Sociology*, 94(1): 25–52.
- ▶ **Hirsch, Barry T.** 2004. “What Do Unions Do for Economic Performance?” *Journal of Labor Research*, 25(3): 415–55.
- Hirsch, Barry T., and David A. MacPherson.** 1993. “Union Membership and Coverage Files from the Current Population Surveys: Note.” *Industrial and Labor Relations Review*, 46(3): 574–78.

- ▶ **John, Philip, and F. R. Whatley.** 1975. "Paracoccus Denitrificans and the Evolutionary Origin of Mitochondria." *Nature*, 254: 495–98.
- ▶ **Jones, Benjamin F., and Benjamin A. Olken.** 2005. "Do Leaders Matter? National Leadership and Growth since World War II." *Quarterly Journal of Economics*, 120(3): 835–64.
- Kremer, Michael, and Benjamin A. Olken.** 2002. "A Biological Model of Unions." Unpublished.
- Lipset, Seymour Martin.** 1960. "The Political Process in Trade Unions." In *Labor and Trade Unionism: An Interdisciplinary Reader*, ed. Walter Galensen and Seymour Lipset, 216–43. New York: Wiley and Sons.
- Lipset, Seymour M.** 1986. "North American Labor Movements: A Comparative Perspective." In *Unions in Transition*, ed. Seymour M. Lipset, 421–52. San Francisco: ICS Press.
- Martin, Donald L.** 1980. *An Ownership Theory of the Trade Union*. Berkeley: University of California Press.
- Michels, Robert.** 1915. *Political Parties*. Translated by Edward Cedar Paul. Glencoe, IL: Free Press.
- Nelson, Richard R., and Sidney G. Winter.** 1982. *An Evolutionary Theory of Economic Change*. Cambridge, MA: Harvard University Press.
- ▶ **Neumann, George R., and Ellen R. Rissman.** 1984. "Where Have All the Union Members Gone?" *Journal of Labor Economics*, 2(2): 175–92.
- Ross, Arthur M.** 1950. *Trade Union Wage Policy*. Berkeley: University of California Press.
- Simon, Herbert A.** 1947. *Administrative Behavior: A Study of Decision-Making Processes in Administrative Organization*. New York: Macmillan.
- Taft, Philip.** 1964. *Organized Labor in American History*. New York: Harper & Row.
- Troy, Leo.** 1990. "Is the U.S. Unique in the Decline of Private Sector Unionism?" *Journal of Labor Research*, 11(2): 111–43.
- Troy, Leo.** 1992. "Convergence in International Unionism: The Case of Canada and the USA." *British Journal of Industrial Relations*, 30(1): 1–43.
- Ulman, Lloyd.** 1955. *The Rise of the National Trade Union*. Cambridge, MA: Harvard University Press.
- Voos, Paula B.** 1983. "Union Organizing: Costs and Benefits." *Industrial and Labor Relations Review*, 36(4): 576–91.
- Voss, Kim.** 1993. *The Making of American Exceptionalism: The Knights of Labor and Class Formation in the Nineteenth Century*. Ithaca: Cornell University Press.
- Wallerstein, Michael.** 1989. "Union Organization in Advanced Industrial Democracies." *American Political Science Review*, 83(2): 481–501.
- Weiler, Paul.** 1983. "Promises to Keep: Securing Workers Rights to Self-Organization under the NLRA." *Harvard Law Review*, 96(8): 1769–1827.
- Weiler, Paul.** 1984. "Striking a New Balance: Freedom of Contract and the Prospects for Union Organization." *Harvard Law Review*, 98(2): 351–420.