Some Economics of Labor Law

Richard A. Posner†

The law governing employment is of vast compass. Among the subjects it embraces are racial and sexual discrimination in employment, the liability of an employer ("master") for the torts of his employees ("servants"), the regulation of occupational health and safety, employees' rights under the pension-regulation law (ERISA), the emerging tort of wrongful discharge of an employee at will, and much else besides. But, to lawyers anyway, the most important subject in the law of employment, as measured by the number of cases, the density of legal doctrine, and other measures of legal activity, remains—even in a period of union decline—the regulation by the National Labor Relations Board of the process by which unions seek to bargain collectively on behalf of workers.1 This regulation is conducted under the authority of the National Labor Relations Act,2 which is the Wagner Act of 1935,3 as amended, principally by the Taft-Hartley Act of 1947.4 When I use the term "labor law" in this paper, I shall, unless otherwise indicated, be referring to this regulatory scheme, even though properly speaking it is just a part of a much larger field.

Whether defined broadly or, as I am doing, narrowly, labor law is as natural a field for the application of economics to law as one could imagine. It regulates explicit markets that have been a subject of continuous and fruitful economic study since Adam Smith’s

† Judge, United States Court of Appeals for the Seventh Circuit; Senior Lecturer, University of Chicago Law School. This is the revised text of a paper given on April 27, 1984, at the Symposium on the Conceptual Foundations of Labor Law, held at the University of Chicago Law School, sponsored jointly with the Social Philosophy and Policy Center of Bowling Green State University. The author is grateful to William Landes, Douglas Leslie, Michael Lindsay, Bernard Meltzer, Melvin Reder, Ronald Schy, George Stigler, and James Talent for many helpful comments on a previous draft of this paper.

1 Regulation by the NLRB is subject to review by the federal courts of appeals, 29 U.S.C. § 160(e) (1982), and on writ of certiorari by the Supreme Court, 28 U.S.C. § 1254 (1982).
day. And though in recent years the focus of labor economics has shifted from unions to other phenomena of labor markets, such as human capital and employment discrimination, there is a rich—and reviving—contemporary literature on the economics of unions. Moreover, as I shall argue in this paper, a well-developed field of economic analysis outside of labor economics—the economic analysis of cartels—can yield to the student of the legal regulation of unionizing many insights.

Yet despite abundant opportunity, there has been relatively little writing in an economic vein about the particulars of labor law, especially—and especially surprisingly—of labor law as I am narrowly defining it. There are, I conjecture (a word used advisedly), two reasons for this situation. The first is that because labor law is doctrinally complex (much more so than antitrust, the economists’ favorite field of law), economists have not found it accessible in the way they have found antitrust law, and more recently

---


* Some exceptions to this generalization should be noted. There is an economically informed literature on the application of the antitrust laws to the union activities that are not exempt from those laws. See, e.g., Leslie, Right to Control: A Study in Secondary Boycotts and Labor Antitrust, 89 Harv. L. Rev. 904 (1976); Meltzer, Labor Unions, Collective Bargaining, and the Antitrust Laws, 32 U. Chi. L. Rev. 659 (1965). There is, of course, an extensive economic literature on the effects of laws regulating wages and hours, industrial health and safety, and employment discrimination. Wrongful discharge is a new area of labor law that has received interesting economic treatment recently. See Epstein, In Defense of the Contract at Will, 51 U. Chi. L. Rev. 947 (1984); Harrison, The “New” Terminable-at-Will Employment Contract: An Interest and Cost Incidence Analysis, 69 Iowa L. Rev. 327 (1984). But economic analyses of specific provisions of the National Labor Relations Act appear to be rare, although I do not pretend to have made a complete search of the literature. I have found a few brief analyses of such provisions by economists. See Jack Hirshleifer, Price Theory and Applications 380-82 (3d ed. 1984); Alchian, Decision Sharing and Expropriable Specific Quasi-Rents: A Theory of First National Maintenance Corporation v. NLRB, 1 S. Ct. Econ. Rev. 235 (1982). Some contributions of economically minded lawyers are cited infra notes 9 & 21. The legal community is not unaware of the economic literature on unions—quite the contrary. See, e.g., Bernard D. Meltzer, Labor Law: Cases, Materials, and Problems 37-94 (2d ed. 1977). But for the most part that literature has not yet been brought to bear on particular provisions of the NLRA.
tort law, accessible. The second reason is that because labor law is (as we shall see) founded on a policy that is the opposite of the policies of competition and economic efficiency that most economists support, the field is unlikely to attract, as a subject for teaching and scholarship, the lawyer who is deeply committed to economic analysis; it is likely to repel him. Of course, you don't have to agree with the normative premises of a field to find it a worthwhile subject for teaching and scholarship. But the fact is—I suppose it reflects the lawyer's training in advocacy—that it is rare for a law professor to make a sustained commitment to a field for whose premises he feels no sympathy at all.

Nevertheless, and somewhat ironically since unions have been in decline in the United States, England, and other countries in recent years, the last few months have seen the appearance of several interesting papers in which economic analysis is brought to bear (in very different ways) on specific problems of labor law in my narrow sense of the term.

One task I have set myself in this paper is simply to make labor law less mysterious to economists, in the hope that they will be encouraged to overcome a natural resistance to immersion in complex legal doctrine. I shall begin therefore with a brief sketch of the American system of labor law and then propose a simple economic model of that system. My basic thesis will be that American labor law is best understood as a device for facilitating, though not to the maximum possible extent, the cartelization of the labor supply by unions. Lest this seem an impolitic (especially for a judge) condemnation of the union movement, I emphasize that I am using the word "cartelization" in a nonpejorative, technical sense: it is the cooperative endeavor of competing sellers to raise the prices of their goods or services (here labor services) above the level that would prevail under conditions of unregulated competition. I take no position on whether it is socially preferable for the price of labor to be determined on a competitive or on a cartelized basis. My analysis is positive, not normative.

I. AMERICAN LABOR LAW

Professor Richard Epstein has conducted a very useful survey of the position of the common law with regard to labor unions. Although that position is typically and not inaccurately described as “anti-union,” Professor Epstein shows that it could just as well be called “pro-competitive,” or, as some economic analysts of the common law would have it, “pro-efficiency.” At common law, labor unions were recognized for what they were: worker cartels designed to raise the price of labor above the competitive level. Picketing, too, was recognized for what it was: an attempt to interfere, by means inherently intimidating, with contractual relationships between the picketed firm and its customers and suppliers, including new workers hired to replace the strikers. So-called “yellow dog” contracts (under which workers agreed not to join unions during the term of their employment) were enforced on the assumption, congenial to classical economic thinking, that the worker was compensated for giving up his right to join a union. If he was not generously compensated, that was nothing to worry about; compensation for not combining with other workers to create a labor monopoly is itself a form of monopoly rent.

It can of course be argued that this picture of an efficient common law of labor relations rests on unrealistic premises about the nature of labor markets, especially in the years prior to the revolution in labor law brought about by the Wagner Act in 1935. If many workers were ignorant of their alternative employment opportunities, wages would frequently have been below the competitive level. If many workers (especially, perhaps, older workers) would have incurred heavy costs by changing jobs, maybe because they had become specialized to a particular employer’s methods or had developed close social and family ties to a particular community or region, employers would have monopsony power, and the workers might be paid less than a competitive wage. If, as Adam Smith believed, conspiracies among employers to depress wages

---

10 See Epstein, Common Law, supra note 9, at 1358-86.
11 This finding provides additional support for the thesis, which I have expounded elsewhere, that the common law is on the whole efficiency-promoting. See, e.g., Richard A. Posner, Economic Analysis of Law 25-191 (2d ed. 1977).
12 See Selig Perlman, A History of Trade Unionism in the United States 147 (1922).
14 See generally Epstein, Common Law, supra note 9, at 1370-75, 1382-85.
15 This description is not wholly accurate; the situation would be one of bilateral monopoly since specialization would also tend to give the workers monopoly power.
were common,\textsuperscript{16} this would be another source of monopsony power. These conditions \textit{may} have been common in the nineteenth and early twentieth centuries in this country, when there were low levels of worker education, a great deal of immigrant labor, a limited number of employers in some markets, no serious enforcement of antitrust laws against employer cartels, and some obstacles to labor mobility (though Americans have always moved around a lot).\textsuperscript{17} But against all this must be set the facts that in the great era of immigration between the Civil War and the end of unrestricted immigration after World War I, America had a chronic labor shortage, which was the main reason for the great immigration; that wages were much higher in the United States than in the rest of the world; and that competition for workers must have been intense and should have limited the extent of monopsony power in labor markets.\textsuperscript{18}

Even assuming that American labor markets were substantially distorted from the competitive norm in ways that unions might have alleviated,\textsuperscript{19} by 1935 these distortions must have been largely in the past (they certainly have a quaint ring today). But whether economically justified or not, the Wagner Act brought about a revolution in the American law of labor relations. The common law was displaced by a system of federal regulation administered by a new agency, the National Labor Relations Board, and designed—as its sponsors and supporters made clear\textsuperscript{20} and as is anyway obvious from the structure of the Act—to foster unionization. In the Taft-Hartley Act in 1947, Congress redressed the Wagner Act’s tilt toward unions somewhat. Legislative and judicial


\textsuperscript{18} Id.

\textsuperscript{19} But not cured: the negotiations between monopolistic unions and monopsonistic employers, a situation of classic bilateral monopoly, will result in fewer employees than under competition because both sides are trying to restrict the supply of labor. Cf. GEORGE J. STIGLER, \textit{The Theory of Price} 207-08 (3d ed. 1966).

innovation since 1947 has greatly expanded the scope of labor law, so that today, as I said at the outset, it extends beyond the regulation of union-organizing activities to embrace the internal governance of unions, racial and other discrimination in labor markets, the regulation of pension plans, and much else besides. But the core of modern labor law remains the NLRB's regulation, under the Wagner Act as amended by the Taft-Hartley Act, of unions' efforts to organize employees and bargain with the employers on their behalf.21

Rather than attempt to summarize the relevant statutory provisions and interpretive doctrines, I will try to convey the essential features of the NLRB's regulation through a description of the process of union organizing and bargaining as it might occur in a small industrial plant.22 The process begins with an employee of a union ("business agent," he is usually called) approaching a friendly employee of the plant (sometimes the plant employee initiates the contact) and giving him union authorization cards to hand out to his fellow employees; when signed, these cards authorize the union to represent the employees who sign them.23 The importance of union authorization cards lies in the fact that if a majority of the workers in the bargaining unit (of which more presently) sign them, the employer may decide to recognize the union as the workers' exclusive representative for collective bar-

---

21 Two other statutes complete the core: the Norris-LaGuardia Act, 29 U.S.C. §§ 101-115 (1982), which among other things greatly restricts the authority of the federal courts to issue injunctions in labor cases, and the Railway Labor Act, 45 U.S.C. §§ 151-188 (1982), which imposes a form of compulsory arbitration on the railroad and airline industries. Compulsory arbitration is also a common legal regime for labor relations in the public sector, which is exempt from the federal labor laws and will not be discussed in this paper, in part because it is already the subject of a rich, and economically well-informed, literature. See, e.g., H. Wellington & R. Winter, The Unions and the Cities (1971); Meltzer & Sunstein, Public Employee Strikes, Executive Discretion, and the Air Traffic Controllers, 50 U. Chi. L. Rev. 731, 738-44 (1983). On the economics of compulsory arbitration, see Ashenfelter & Bloom, Models of Arbitrator Behavior: Theory and Evidence, 74 Am. Econ. Rev. 111 (1984).

22 The reader who wants greater detail and citations to cases is advised to begin with Robert A. Gorman, Basic Text on Labor Law: Unionization and Collective Bargaining (1976). This is a lucid, compact, and relatively nontechnical introduction to the field. No extensive knowledge of law is required to be able to read it with understanding and profit. Also very good and more up-to-date, though longer, is the two-volume The Developing Labor Law (C. Morris 2d ed. 1983). For a brief, serviceable description of the federal labor statutes for nonlawyers, see F. Marshall, A. King & V. Briggs, supra note 5, at 426-52.

23 See R. Gorman, supra note 22, at 41. The reason the business agent will work through one or more plant employees, rather than distribute the cards himself, is that the Board allows the employer to forbid union solicitation on his premises. The Board's position rests on the practical ground that a stranger's presence on the premises can disrupt work discipline and in some cases can be a hazard to the employees' safety.
gaining without the formality of a representation election. More important, if at least thirty percent of the workers sign authorization cards and the employer refuses to recognize the union, the Board will order a representation election.

The efforts of an employee to induce his fellows to sign union authorization cards would often, in the absence of legal protection or of successful concealment by the employee of his activities, be set at naught by the employer's firing him. This would be an example of rational predatory action. It is true that the employer would impose a cost on himself by firing the worker, assuming that he was a satisfactory worker (and if he were not, he probably would have been fired already). But the cost would be small compared with the benefit to the employer of signaling to the remaining employees that if any one of them stepped forward to take the place of the fired employee as the union's organizer, he would be fired too. True, if the workers hung together and struck in support of the fired employee, the balance of costs would be altered and the employer might back down. But since the workers would be unorganized (for I am speaking of how an employer might try to thwart an organizational drive), a strike might be difficult to arrange: the workers would face classic free-rider problems. Those problems, however, should not be exaggerated. There were independent unions (as well as "company unions," which the Wagner Act forbade) long before the Wagner Act was passed. But the fraction of workers who were unionized rose very rapidly after the Act was passed, and this is some evidence that it was indeed difficult to organize workers without the protections that the Act extended to union-organizing efforts.

The key protections are in the sections of the Act that entitle employees to engage in concerted activities and that make it unlawful for the employer to interfere with those activities. Firing an employee because he is trying to organize the plant presents a

---

24 See id. at 230.
26 For an alternative characterization, see infra notes 52-55 and accompanying text.
28 For an interesting, if dated, treatment of independent unions, see S. Perlman, supra note 12. Incidentally, chapter 7 contains some interesting discussion of common law attitudes toward labor unions.
29 See National Labor Relations Act §§ 7-8, 29 U.S.C. §§ 157-58 (1982); see also Inter-Collegiate Press v. NLRB, 486 F.2d 837, 845 (8th Cir. 1973) ("Conduct having even a 'comparatively slight' impact on employee rights may be a violation of § 8(a)(3), unless the employer has established a legitimate and substantial business justification." (citation omitted)).
clear case of unlawful interference, as do much milder forms of retaliation—even something as trivial as not inviting the employee to a company party. The employer is thus denied the natural advantage that he would have, as one facing many, in fending off organizing activities. In addition, "yellow dog" contracts are forbidden by section 3 of the Norris-LaGuardia Act.

Let us assume that the union organizer has gotten signatures from thirty percent of the employees. The next step chronologically is the election campaign, but before getting to that I must pause briefly to discuss the electoral unit, or the "bargaining unit" as it is called. It is not a synonym for the firm, or even for the plant. Rather, it is any group of employees that the Board decides is sufficiently homogeneous, and sufficiently distinct from other employees, to be allowed to form its own bargaining unit. Ordinarily, though not always, the unit will be limited to one plant even if the firm owns other plants as well. Often there will be more than one unit in the plant or facility. For example, a single hospital, whether or not part of a chain, might contain separate units for doctors, for registered nurses, for nurses' aides and other maintenance employees, and perhaps for technical employees such as X-ray technicians. The Board's discretion in determining the appropriate bargaining unit for a particular type of firm is broad, but there are some restrictions on it; most important, the Taft-Hartley Act denies protected status to supervisory employees, from foremen on up, unless their supervisory responsibilities are incidental (e.g., a doctor supervising his secretary).

Only one question is put to the electorate—the members of the bargaining unit—in the representation election: whether to make the union that is trying to organize the unit the exclusive agent of the unit employees for purposes of bargaining with the employer over wages and working conditions. The outcome of the election is determined by majority vote of the employees in the unit, voting by secret ballot. The election is preceded by a campaign that in some ways is like a political campaign. But it is shorter, and the voting is on whether to unionize rather than on candidates for office. Furthermore, the contending parties—union

---

* See NLRB v. Village IX, Inc., 723 F.2d 1360, 1366-67 (7th Cir. 1983).
* See National Labor Relations Act § 9(b), 29 U.S.C § 159(b) (1982).
* See id. § 9(a), 29 U.S.C. § 159(a) (1982).
and employer—are more limited in what they are allowed to say than are candidates and supporters in political elections: not only must the employer refrain from firing union adherents or otherwise interfering with the union's campaign, but he may not threaten retaliation if the union wins or promise specific benefits if the union loses; promises of benefits if the candidate wins are of course a staple of true political campaigns.

If the union loses a valid representation election, the Board will not direct another election for a year, and then only if the union again gets at least thirty percent of the employees to sign union authorization cards. If the union wins the election the consequences are more complicated. First, all the employees in the bargaining unit, whether or not they voted for the union and whether or not they want to belong to it, are forbidden to bargain individually with the employer; the union is as much the exclusive bargaining representative of the dissenters as of the employees who voted for it. Second, all the employees, again regardless of their personal sympathies, must, if the collective-bargaining agreement between the employer and the union so provides (and it is a provision for which unions press very hard in negotiations), pay union dues and often must actually join the union. Third, the employer must negotiate with the union in good faith for a collective-bargaining agreement that will specify the terms and conditions of employment of the members of the unit for a specified period, usually one to three years.

But the employer is not required to yield to the union's demands even in part (which makes one wonder whether the duty to bargain in good faith has much bite), and often he will not. In that event the union may decide to call a strike in an effort to win at least partial agreement to its demands. If it does not call a strike, even though the employer has made no significant concessions to its demands, the union may lose the workers' support: they will see that they are getting nothing in exchange for union dues that are

---

36 See NLRB v. Exchange Parts Co., 375 U.S. 405, 409 (1964) ("We have no doubt that [the NLRA] prohibits not only intrusive threats and promises but also conduct immediately favorable to employees which is undertaken with the express purpose of impinging upon their freedom of choice for or against unionization and is reasonably calculated to have that effect.").


38 See 29 C.F.R. § 101.18(a) (1983).


not trivial.

The economic function of the strike requires consideration at this point. It is related to the bilateral-monopoly character of labor-management negotiations. When a nonlabor market becomes cartelized, members of the cartel raise their price and, anticipating some substitution away from their product by consumers, reduce output, but not to zero. But if there were only one consumer for the cartel’s product, he might say to the cartel, "I won’t buy from you at the higher price," and they would then face the choice of either backing down or not selling to him. This happens occasionally in nonlabor markets, but in labor markets it happens often. The union deals with a single employer (or several employers bargaining as one in a multi-employer bargaining unit), who may be tempted to refuse to accept the union’s demands (i.e., may threaten to buy nothing rather than come to terms), and then the union must either strike in order to enforce its terms or else back down. The union cannot just write off this “customer” as marginal, as a product monopolist often can when he raises his price; for each employer’s work force will be represented by its own local union (often more than one), and if the union ignores the workers’ interests they will vote the union out and the employer will be free to go his own way. Thus we have a classic example of bilateral monopoly: the union and employer can deal only with each other and a refusal to deal, by imposing costs on the other party, makes him more likely to come to terms. The strike imposes costs on both parties: on the employer, by forcing him to reduce or cease production, and on the workers, by stopping their wages. The balance of those costs will determine the ultimate settling point between the union’s initial demand and the employer’s initial offer.

Labor law affects these costs. For example, the Board allows the employer, if there is a strike, to hire replacements for the striking workers. He is even allowed to offer the replacement workers permanent jobs—and to do so even if such an offer would not be necessary to induce them to work for him. It would never be necessary if the employer were permitted to pay a wage high enough to induce a replacement to work temporarily, without promise of a permanent job. But the employer is not permitted to pay replacement workers a higher wage than he paid the workers who have struck. This rule shifts the balance the other way; it limits the em-
ployer's ability to hire replacements, permanent or temporary.

Although, subject to this qualification, the employer may hire permanent replacements, he may not fire the striking workers who have been replaced. \(^4\) True, unless the strike was provoked by an employer's unfair labor practice, the employer does not have to reinstate all of the strikers as soon as the strike ends or pay any of them their back wages. But when the strike is over, those strikers whose places have been filled by permanent replacements must be put at the head of the queue, to be reinstated as vacancies appear, and those strikers whose places have not been filled must be reinstated immediately. \(^4\)

Attempts to defeat strikes by hiring replacement workers are less common than one might expect; more common is the use of supervisory personnel to replace the striking workers temporarily (hence the importance of the National Labor Relations Act's exclusion of such personnel from the Act's protections). The problem with using replacement workers is that in order to get to the workplace they will have to cross the picket line thrown up by the striking workers' union. Even though picketers are not legally privileged to use force to prevent the crossing of picket lines, whether by replacement workers or by customers or suppliers of the picketed establishment, there is often a latent threat of violence (which cannot, however, be used as a ground for firing or enjoining a picketer \(^4\)), especially against replacement workers ("scabs"). And in pro-union communities the police may not have the desire or ability to control this threat effectively (though they may come down hard on any effort by the employer to hire "goons" to intimidate the picketers). Usually the picketing workers can at the very least identify the replacement workers, who may therefore fear eventual retaliation even if the picketing itself is completely peaceful. Their fear will be enhanced by the Act's provision forbidding the employer to fire striking workers. When the strikers eventually return to work, they will be working side-by-side with the permanent replacements, who may entertain fears for their own safety or at least for the continued congeniality of the workplace.

\(^{43}\) 29 U.S.C. § 152(3) (1982) preserves the strikers' status as "employees" protected by the NLRA. See also NLRB v. Fleetwood Trailer Co., 389 U.S. 375, 378 (1967) ("[U]nless the employer who refuses to reinstate strikers can show that his action was due to 'legitimate and substantial business justifications,' he is guilty of an unfair labor practice." (citation omitted)).

\(^{44}\) See NLRB v. Fleetwood Trailer Co., 389 U.S. 375, 379 (1967).

\(^{45}\) See, e.g., Chevron U.S.A., Inc. v. NLRB, 672 F.2d 359, 380-61 (3d Cir. 1982); NLRB v. W.C. McQuaide, Inc., 552 F.2d 519, 527-28 (3d Cir. 1977).
If a collective-bargaining contract between union and employer is signed, with or without a strike, it will be judicially enforceable in accordance with a federal common law of collective-bargaining contracts.\textsuperscript{46} Often such contracts contain no-strike clauses, and if such a clause is violated, the employer may be able to get an injunction against the strike and an award of damages against the union.\textsuperscript{47} Whether or not there is a no-strike clause, a "wildcat" strike—a strike not authorized by the union—is not protected activity if it has a tendency to interfere with the union's role as exclusive bargaining representative;\textsuperscript{48} and if a strike is unprotected, the employer can fire the wildcat strikers with impunity.

Unlike an elected public official, a union that is elected to be the collective-bargaining representative of some unit does not serve a fixed term. But upon a showing that the union probably has lost majority support the employer can file an election petition or can refuse to bargain with the union and thus force the union to file such a petition.\textsuperscript{49} In such a case the Board will order a new election if at least one year has elapsed since the union was certified as the unit's bargaining representative.\textsuperscript{50}

\section*{II. UNIONS AS LABOR CARTELS}

Cognoscenti of labor law will recognize the preceding discussion as but a crude thumbnail sketch of the law of collective bargaining. But it will serve to frame an inquiry into the economic logic of that law. My discussion will be illustrative rather than exhaustive: multi-employer bargaining, secondary boycotts, and antitrust restrictions on union activity are among the relevant topics that I have omitted in the interests of time and space.

If unionization is a means of cartelizing labor markets, the National Labor Relations Act, which even with the Taft-Hartley amendments plainly fosters unionization, is likewise a means to cartelize such markets. Economists have long treated unions as labor cartels,\textsuperscript{51} though alternative explanations have been ad-

\begin{itemize}
\item Boys Mkts., Inc. v. Retail Clerks' Local 770, 398 U.S. 235, 252-54 (1970).
\item See Emporium Capwell Co. v. Western Addition Community Org., 420 U.S. 50, 70-73 (1975) (unauthorized strike by minority employees to protest discrimination not protected by NLRA).
\item See, e.g., J. HIRSCHLEIFER, supra note 8, at 380-82; G. STIGLER, supra note 19, at 268-70; Friedman, Some Comments on the Significance of Labor Unions for Economic Policy, in THE IMPACT OF THE UNION 204 (D. Wright ed. 1951); Lazear, A Microeconomic Theory of
\end{itemize}
One is that the way in which unions benefit their members is not by reducing the supply of labor (and hence forcing up the price, i.e., wages), but by increasing the productivity of the work force. This they are said to do in various ways. One is by providing a vehicle for collecting, and communicating to the employer, workers' complaints about wages and working conditions.

In the absence of such a vehicle, it is argued, workers might be afraid to voice their complaints, and the employer would learn of them only indirectly and belatedly, by observing a higher quit rate. Another example: unions invariably press for inclusion, of a provision forbidding management to fire workers except for good cause, and requiring it, when it lays off workers because of an economic downturn, to lay them off in reverse order of seniority (i.e., juniors first). When such job security is lacking, as is usually the case in nonunion firms, the older, more experienced workers may—it is argued—be reluctant to share their know-how with the younger, newer employees, fearing that if they do the younger employees will then be competing for their jobs. As a result of this reluctance, productivity is thought to suffer.

Although some empirical support has been marshaled for this productivity-enhancement theory of unionization, the theory is extremely hard to accept. It is inconsistent with the fundamental assumption of economics: that people, in this case employers, are rational profit or utility maximizers. Although this assumption may not hold true in all settings, the behavior of business employers towards their employees is one setting where it probably does.

---


53 See, e.g., Brown & Medoff, supra note 52, at 356-59; Freeman, supra note 52, at 365.

54 See, e.g., Freeman, supra note 52, at 366 (unionism is a "market mechanism for imparting information, aggregating preferences, [and] altering authority relations"); see also Freeman & Medoff, supra note 52, at 70-74.

55 See Brown & Medoff, supra note 52, at 362-69; Freeman & Medoff, supra note 52, at 78-87.
If granting his employees tenure will increase their productivity, the rational employer will do so, for this will reduce his costs of production. Even if the whole productivity gain is paid to the employee in the form of a higher wage, the employer will be better off. He will have lower total costs than his competitors and will therefore be able to expand his output relative to theirs and increase his profits. Even if only a single employer in a competitive industry tumbled to the advantages of granting tenure, competition would force the others to follow suit. And so with encouraging workers to complain rather than waiting for them to quit: the rational employer will encourage them to complain, by cash rewards or whatever it takes, if worker turnover is costly to him.

The proposition that unions enhance productivity also flies in the face of massive, if unsystematic, evidence pointing to the opposite conclusion. Featherbedding seems a more common attribute of unionized than of nonunionized work forces (at least in the private sector); many industries that are heavily unionized are notable for their low productivity; and for every older worker whom job security encourages to share his know-how, casual observation suggests that there is at least one other older worker, and probably several, whom job security protects at the expense of a more efficient younger worker. Most important of all, for many generations now employers have expended substantial resources to prevent unionization of their plants—expenditures that would be irrational if it were true that unions enhanced labor productivity. Such persistent irrationality by American businessmen is very hard to credit, but it is a proposition entailed by the productivity-enhancement theory of unionization.

It seems far more plausible to assume that the intended and actual effect of unionization is to raise the price of labor above the competitive level, and to depress the supply of labor below the competitive level, in the unionized sector (about twenty percent of the American work force is unionized). This view not only is commonsensical but explains a wide range of phenomena. It explains the support of unions for the minimum wage, which has the effect of raising the price of substitute nonunion labor, and for government regulation of workplace safety, which reduces competition from nonunion employers. It also explains the pattern of unioniza-

---

56 These points are neglected by Freeman & Medoff, supra note 52, at 91-93, in their attempt to explain management opposition to independent unions.

tion in the American economy, which is about what one would predict from differences in the ability to cartelize the labor supply in different industries. Thus we predict that we will find, and do find, the most effective unions in industries where competition among employers is weak (often because of government regulation), the cost of the organized work force is a small part of the employer's total costs, and the employer produces a nonstorable commodity, so that a strike will impose heavy costs on him. Excellent examples of all three factors (all of which are different aspects of labor-supply inelasticity) are found in the airline pilots' union before the deregulation of the airline industry and in the railroad industry in its heyday, where unionization took hold long before government came directly to its aid. Finally, as we will now see, the cartel theory of unionization explains better than any alternative theory the dominant features of the regulation of labor relations by the National Labor Relations Board.

The theory of cartels teaches that cartelization of a market is a very difficult, perhaps hopeless, endeavor if there are a large number of competitors. And that is the typical situation in labor markets. It is not only that the work force of all but the smallest employers will contain far more members than has been thought the limit for effective cartelization without government assistance (a critical qualification in the present context, obviously); in addition, the relevant market includes workers employed by other firms (or unemployed) who, for a slightly higher wage, would go to work for an employer facing a strike.

These workers are an important part of the relevant market. In the theory of cartels, potential entrants are important only when the number of firms actually selling in the market—a number corresponding in the labor market to the number of employees actually selling their services to the employer in question—is small. If the number of significant firms is large (the qualification being added to exclude the case where a few firms have most of the sales and there is an unimportant fringe of tiny firms), cartelization probably will fail because each firm can expand its output and will be irresistibly tempted to do so if others reduce their output. (If none could expand its output, then a reduction in output by even a single firm would push the market price above the competi-
tive level because the market's total output would be smaller as a result of that reduction.) Now it is easier for a firm to expand its output than for an individual worker to do so. The firm can add to its work force or to its capital; the individual worker would have to work harder or work longer hours. Of course this is possible within limits, especially for a short time. And a short time may sometimes be good enough: since a strike is costly to the striking workers, keeping the firm operating for a short time may be sufficient to break the strike even if the firm is forced to contract its operations—provided it is not forced to shut down completely. But if the strikers have more staying power than this, their strike may be effective though far fewer than all the workers join it, for the remaining workers may not be able to take up the slack by working harder, or for longer hours, for as long as it would take to break the strike. The strike might last too long for nonstriking workers or supervisors to be able to keep the plant operating and too long for the employer to substitute capital inputs for the labor inputs no longer available to it. In either case the firm's ability to hire replacement workers from other employers or from the pool of unemployed workers could determine the success or failure of the strike.

The large number of potential competitors of the striking workers is such a large obstacle to cartelizing labor markets without governmental assistance that most union-organizing efforts probably would be ineffectual without such assistance, provided the government enforced against unions as against the rest of society the basic laws protecting rights of property, contract, and personal safety (so that unions could not use force or the threat of force to achieve their ends). We now have to consider how the National Labor Relations Act alleviates the large-number problem and in other ways fosters effective if incomplete cartelization of labor markets.

To begin with, through the concept of the employer unfair labor practice, the Act prevents the employer from engaging in the kind of rational predatory activity that, as I suggested earlier, could be used to defeat unionization in its incipient stage. Put differently (for those skeptical of the economic rationality of predatory behavior in any form), the Act prevents competition between two groups of workers: those willing to work for the competitive wage and those willing to devote time to (and take risks in the hope of) obtaining a higher wage through unionization. The employer is forbidden to substitute members of the former group for members of the latter; it is as if a consumer were forbidden to
switch his patronage to price cutters.

Next, the Act increases the wealth of unions and thus helps them play their vital role as agents for organizing workers. The union's role corresponds to that of trade associations, exclusive sales agencies, the old railroad rate bureaus, and other institutions for organizing competitors in product markets, but the union is more essential because of the large number of competitors to be organized. The Act, as interpreted by the Board and the courts, helps unions in several ways. It forbids the employer during the union-organizing campaign to offer (or even promise) its workers the higher wages or better fringe benefits that the union has promised to press for. Such an offer, if accepted, would undermine the union by preventing it from recouping the expenses of organizing by collecting union dues. The Act protects unions from another form of free riding by forbidding workers, after the union has been certified as the exclusive bargaining representative, to negotiate separately with the employer and by empowering the union, without regard to the wishes of individual members, to negotiate a provision in the collective-bargaining contract requiring all members of the bargaining unit to pay union dues. Such a provision prevents an individual worker from obtaining the benefits of unionization without paying his share of the costs. Without dues, unions could not function. Indeed, assuming that what unions seek to maximize is their dues income, if there is competition between unions that income will be proportionate to the benefits that the union confers on the workers it represents. The union's income would in any case be much less if a worker could enjoy the benefits conferred by the union without paying any dues.

The devices for preventing free riding on a union's organizing and other activities are very far from being perfect. If an employer, in an effort to discourage a union from organizing his workers, pays a wage that is less than the union scale by a smaller margin than the union's dues—as he can do without violating the Act—both the workers and the employer will be better off than if the union or-

---

69 The Taft-Hartley Act, however, allows the states to forbid "union security clauses," as they are called, see National Labor Relations Act (as amended by Taft-Hartley Act) § 14(b), 29 U.S.C. § 164(b) (1982), and a number of states, disproportionately southern, have taken up this option, see, e.g., Ala. Code §§ 25-7-30 to -36 (1975); Ga. Code Ann. § 34-6-21 to -28 (1982); Miss. Code Ann. § 71-1-47 (1972).

60 There is great debate over just what it is that unions maximize. For a discussion of contending positions, see Donald L. Martin, An Ownership Theory of the Trade Union 6-30 (1980). Dues maximization seems the natural assumption but is not essential to my analysis.
ganizes the workers. Yet it is only the threat of unionization that enables this benefit to be obtained, and the union receives no compensation for creating it. Furthermore, although every worker must pay union dues once the union has become the collective-bargaining agent for his unit and has negotiated a union security clause with the employer, the union cannot force the workers to honor a strike call (unless they are union members—not just dues-payers—and have not quit the union before crossing the picket line).

Much like the fringe firm in a cartelized market, the individual worker may seek the best of both worlds by continuing to work during the strike while hoping that the union will succeed in wresting concessions from the employer so that after the strike the worker’s wages will be higher as a result of it. If enough workers think this way, the strike will fail and all the workers may be worse off than if they had joined it. But this is the same phenomenon as occurs when a cartel of product sellers fails because of defections by members of the cartel who think they can have the best of all worlds by free riding. Such failures are common.

What limits the form of free riding that consists of refusing to honor a strike call is a practical sanction that has no counterpart in nonlabor markets. The worker who continues to work during the strike knows that once it is over he will be working side-by-side with the workers who struck (unless all of their places are filled by permanent replacements), and he may fear retaliation in forms difficult to detect and prevent. Even if the strikers have been permanently replaced, the workers who refused to honor the strike will know that the strikers may eventually come back to work because, as noted earlier, the Act puts the strikers at the head of the queue to be hired (technically, reinstated) when vacancies occur. The prospect of eventually finding oneself working side-by-side with the former strikers will not only increase the likelihood that a strike call will be honored by all; it will also, as I mentioned earlier, discourage some new workers from signing on as permanent replacements in the first place, especially since they cannot be paid a higher wage for doing so.

Genuinely peaceful picketing is thus the counterpart in the la-

---


42 There is divided authority on a union’s right to prevent an employee from resigning from the union during a strike. Compare Pattern Makers’ League v. NLRB, 724 F.2d 57 (7th Cir. 1983) (allowing resignation), cert. granted, 53 U.S.L.W. 3235 (U.S. Oct. 1, 1984), and International Assoc. of Machinists, Local 1414, 270 N.L.R.B. Dec. No. 209 (June 22, 1984) (same), with Local 1327, Int’l Ass’n of Machinists v. NLRB, 725 F.2d 1212 (9th Cir. 1984) (prohibiting resignation).
bor setting of the practice (required, for example, in the rail and trucking industries by the Interstate Commerce Act) of pricing in accordance with published tariffs. The published tariff shores up a cartel by enabling competitors to detect cheating on the cartel price immediately. Picketing serves a similar function by enabling the striking workers, corresponding to the members of a cartel who observe the cartel price, to identify any member of the cartel (i.e., any fellow worker) who is cheating by continuing to work during the strike. In this analysis, picketing is not really an informative activity (setting aside the information that is implicit in any threat); it is an information-gathering activity.63

The cartel analogy may help explain why unions invariably insist that the collective-bargaining contract provide some form of job security. No doubt, part of the reason is merely to back up the law's prohibition of discrimination against union supporters,64 but the theory of cartels suggests a further point. An important object of job-security provisions is to obtain preferential treatment for senior workers. Some workers laid off during a business downturn will find other jobs during the period of layoff and not return to their original employer, who will therefore be hiring replacements for them. And just by the workings of chance, these replacements may be less well disposed to the union than those who were laid off and later quit. So the union will want some criterion for the order of layoffs that will ensure so far as possible that those workers who are least likely to favor the union will be laid off first. These are the younger workers.

Much casual observation supports this proposition, but it also has a theoretical basis. Younger workers are more mobile than older ones. The older ones are more likely to have family obligations that make it difficult to relocate geographically, and their

63 This has possible implications for the analysis of the first amendment rights of pickets, but I shall not attempt to develop those implications here.

64 Besides overt discrimination, employers might find subtle ways of discouraging unionization. For example, workers must differ in their propensity to vote for unions, to go out on strike, and otherwise to engage in cartel-promoting behavior. Therefore, in the absence of contractual job protection, the employer, after discovering that a majority of his workers wanted a union, might discharge some of the workers at random. (I am now assuming that he would not try to discharge solely, or disproportionately, those whom he knew to be union adherents, because that would be clearly unlawful conduct.) His hope would be that the replacement workers might, simply by chance, contain a lower proportion of union supporters, so that he might eventually be able to get the union decertified. Of course this would be a sensible strategy only if the employer thought that union support among his existing work force was above average for his industry, location, etc. The strategy would violate the law, but would be more difficult to detect than the firing of just (or mainly) union supporters.
human capital may have become specialized to the particular job they are doing for their employer (assuming that the older worker, on average, has worked longer for this employer than has the younger worker). Many younger workers are temporary employees, trying out one job after another; some are teenagers working part-time and bound for very different careers. Being less mobile, the older workers are more at the mercy of the employer (like shareholders whose shares are not freely tradable) and therefore have more to gain even in the short run from unionization. They also are more likely to be around to enjoy the benefits that the union generates for the workers in exchange for dues (the collection of dues begins before any of those benefits are realized). True, the younger workers, if they do stick around, will enjoy those benefits longer. But the discount rate applied to benefits from unionization other than those that can be realized in the immediate future must be high, not because workers are short-sighted, but because the union may be decertified or the plant closed before the benefits are realized. An additional point is that, at least in jobs that require strength or stamina, older workers may be less productive than younger workers, with whom—but for union-negotiated seniority protection—the older workers would be competing.

If this analysis is right, then by requiring the younger workers to be laid off first, the union is less likely to lose union adherents than if layoffs were random with respect to age. Moreover, they would never be random. The employer not confined by a collective-bargaining agreement would want to lay off the least productive workers first. They are likely to be disproportionately older and in any event disproportionately pro-union, for it is the least productive employees (whatever the reason why they are least productive) who fare the worst if wages are determined on a competitive basis.

This analysis also explains why unions want employers to use seniority to determine the order of layoffs even though productivity might be maximized, to the mutual benefit of employer and employees, if the union allowed the employer to choose whom to lay off in return for the generous compensation of any older worker laid off. Even if senior workers were made whole, there would still be a disadvantage from the union’s standpoint: some of those laid off would find other jobs and therefore not return to their original employer when the layoff ended, and they would be replaced by younger workers less likely to support the union. Finally, we should note that a seniority rule, by making the employer’s work force less mobile (senior workers have more to lose from quitting), generates additional support for the union.
Another important factor facilitating or retarding the organization of a plant or other facility is the determination of the bargaining unit (the electoral unit for the representation election). In general, the larger the unit the better off the employer is, and the smaller the unit the better off the union is. The larger the unit is—that is, the more employees it has—the more difficult it will be for the union to obtain the majority vote that it needs in order to be designated the exclusive bargaining representative for the unit. This is not only because it takes more resources in absolute terms to get more votes (a national political election is more costly than a local one), but also because the members of the unit are more likely to have divergent interests with respect to tradeoffs among wages, fringe benefits, job security, and workplace safety. This will make it difficult for the union to appeal to a majority and, even if it gets a majority, will make it difficult for the union to formulate a coherent set of demands and enforce those demands by an effective strike threat. This is much like the problem of fixing prices in a producers' cartel when the producers have dissimilar cost functions.

A potentially offsetting factor is that a strike by a small unit may not impose substantial costs on the employer, in which event the union and the workers will gain little (in dues and in wages, respectively) from a successful organizing campaign, even if it is cheap to conduct. But if the unit is small precisely because the workers who comprise it do a different type of work from the other workers in the plant (so that making them a part of a larger unit would result in a heterogeneous unit), it is quite possible that if they go out on strike the plant will have to close down; the work they do, not being duplicated elsewhere in the plant, may well be essential. In addition, a small unit may be large relative to the size of the plant or facility in question. Both points are illustrated by health-care facilities (mainly hospitals and nursing homes), where unions have made great strides since the NLRB's authority was extended to nonprofit health-care facilities in 1974. A hospital may have a small number of employees overall, divided as I noted earlier into several units (doctors, registered nurses, etc.), and a strike by any unit might close the facility down. Since the employer cannot produce for inventory, it will incur very substantial costs from even a short strike. This is why the law requires that

---

66 See R. Gorman, supra note 22, at 67-68.
unions give ten days' notice of a strike in a health-care facility;\(^6\) it is another example of how current law tempers the pro-union policy introduced by the Wagner Act.

Professor Douglas Leslie has suggested that unions would often be better off with larger units because this would facilitate the mediation of conflicts among subgroups of employees.\(^8\) If you have three local unions in a plant, however, their presidents should be able to negotiate some arrangement for mutual support; it is a negotiation among just three people, which the Coase Theorem suggests should be feasible, though there are possible "trilateral monopoly" problems and additional complications stemming from the fact that they will be negotiating in a representative capacity. But if the negotiation is within a unit, no faction has a representative who can negotiate on its behalf; the costs of negotiation will therefore be (I should think) higher; and if so the probability of unresolved conflict will also be higher. I am therefore led to predict that in periods when the NLRB is dominated by Democrats (whom most union leaders support), the Board will tend to certify smaller bargaining units than in periods when Republicans dominate. This would be a fruitful subject for empirical research.

If I am right in my contention that the National Labor Relations Act is best understood as a means of federal governmental support for the cartelization of the labor supply, this may also illuminate another feature of the Act: the vesting of primary responsibility for enforcing it in an administrative agency, the NLRB, rather than in the courts. Since the Act turned labor policy on its head, transforming a public policy of fostering competitive determination of wages and working conditions into one of fostering cartelization, it was quite sensible for Congress to be concerned that state and federal judges—who after all had largely fashioned the former policy—might resist its inversion. It would have made less sense if all the Act were doing was enhancing labor productivity—though Congress might have feared that the judges would misunderstand that this is what the Act was doing.

All that was years ago, and now there are very few judges, state or federal, who have any emotional or intellectual commitment to competitive labor markets. Although the word "cartelization" has negative overtones (more so, indeed, than in the 1930's, when the Depression was attributed in some quarters to excessive

---

\(^{6}\) See National Labor Relations Act (as amended by Health Care Institutions Amendments Act) § 8(g), 29 U.S.C. § 158(g) (1982).

\(^{8}\) Leslie, supra note 9, at 50.
competition), I am sure that most judges today would agree that if federal labor policy is one of facilitating the cartelization of labor, they should, and without much pain can, use this policy to guide them in reviewing the decisions of the NLRB. The only real difficulty is that with the Taft-Hartley amendments, the National Labor Relations Act no longer evinces a univocal policy of promoting cartelization. Even in its pristine Wagner Act form, the NLRA did not totally embrace such a policy. For example, the Act has since the early days been interpreted to allow employers to replace strikers, and has also been interpreted not to protect concerted activity that involves a danger of physical destruction (e.g., damaging the employer's machinery) or personal injury. The rationale of this exception is not quite so obvious to an economist as it might appear to be. Strikes that destroy much more valuable intangible assets are protected. But there is a difference, and the exception for destruction of tangible assets does limit the power of unions. Destroying intangible assets (business goodwill, customers' time, etc.) usually requires a lengthy strike, which is costly to the workers as well as to the employer, his customers, and his suppliers; equally costly destruction of tangible assets might be accomplished in minutes.

A more ambiguous example of a limitation on the union-promoting policy of the Act (as it has been interpreted) is the requirement that the union get at least thirty percent of the workers in the bargaining unit to sign union authorization cards before a representation election will be ordered. It is not obvious that lowering the threshold would promote unionization. A weak union might get enough signatures to compel an election, then lose it resoundingly and by doing so make it harder for a stronger union to organize the plant subsequently.

But the Taft-Hartley Act did make a difference. Notably, by withholding the protection of federal law from supervisor unions (and as a result there are few such unions and most are powerless), the Act strengthened the hand of employers by enabling them to substitute for strikers other workers less likely than permanent replacements to be intimidated by returning strikers. It also outlawed the closed shop, which is a device that minimizes free rid-

---

69 See NLRB v. Fansteel Metallurgical Corp., 306 U.S. 240, 255 (1939) ("We are unable to conclude that Congress intended to . . . invest those who go on strike with an immunity from discharge for acts of trespass or violence against the employer's property . . . ").

ing on union efforts by requiring the employer to hire from the ranks of those who already belong to the union, thus excluding those who join after the plant has been organized.

But the impact of the Taft-Hartley Act is easily exaggerated, as another example will show. Although the Act made no-strike clauses enforceable by damage suits against unions, it is very hard to see this provision as anti-union. A union doesn’t have to agree to such a clause; and if it does, presumably it has been compensated for it. Expanding freedom of contract ought to benefit all parties to a potential transaction. It would be different if the Act allowed “yellow dog” contracts. Those are not contracts between unions and employers but between individual workers and employers and are a device by which employers can exploit the large-numbers problem that complicates unions’ organizing efforts. Each worker knows that his signing an agreement with his employer not to strike while he is employed will have little effect on the success of any union organizing efforts in his plant because he is one of many; knowing this, he will sign such an agreement for only a modest consideration. If all or at least most workers think the same way (and why shouldn’t they?), the employer will have succeeded in preventing union organizing at his plant for a total cost that may be much less than he would have to pay in higher wages if the plant were organized (provided there is not already in being a strong union that can pay the workers more than the company can pay to induce them not to sign “yellow dog” contracts). The banning of “yellow dog” contracts (accomplished in the Norris-La-Guardia Act a few years before the Wagner Act) not only is a rational component of a labor policy dedicated to facilitating labor cartels but is perfectly consistent with the provision in the Taft-Hartley Act allowing no-strike clauses to be enforced. Indeed, the federal labor laws as a whole appear to have a remarkable consistency and intelligibility when viewed as a legal regime for fostering (though not to the maximum possible extent) the cartelization of labor markets.