The Historical Origins and Economic Structures of Workers' Compensation Law

Richard A. Epstein

Follow this and additional works at: https://chicagounbound.uchicago.edu/journal_articles

Part of the Law Commons

Recommended Citation

This Article is brought to you for free and open access by the Faculty Scholarship at Chicago Unbound. It has been accepted for inclusion in Journal Articles by an authorized administrator of Chicago Unbound. For more information, please contact unbound@law.uchicago.edu.
I. General Considerations

THE HISTORICAL ORIGINS AND ECONOMIC STRUCTURE OF WORKERS' COMPENSATION LAW*

RICHARD A. EPSTEIN**

During the nineteenth century, the nature and extent of an employer's liability for industrial accidents were the object of intense debate, both within and beyond the legal arena.¹ Since negligence was widely regarded as the proper basis of liability, the require-

---

* The John A. Sibley Lecture in Law delivered at the University of Georgia School of Law on April 6, 1982, revised and expanded for publication.

** James Parker Hall Professor of Law, University of Chicago. I would like to thank Ross Green and Mark Holmes for their helpful research assistance in the preparation of this Article.

¹ See, e.g., Thompson, Under What Circumstances A Servant Accepts the Risk of His Employment, 31 AM. L. Rev. 82, 85-86 (1897), which contains this denunciation of the established order:

I do not want my professional brethren to think for one moment that I balance the life of a railway brakeman against the slight expense to a railway company of blocking its frogs and switches. I should be sorry to have them think that I ever was willing to balance the life of a railway brakeman against the slight expense to a railway company of building the upper works of its bridges sufficiently high for a brakeman to stand upon the top of his car without coming in contact with them. These are murder-machines; and the rule of judge-made law which holds the servant at all times and under all circumstances, bound to avoid them at his peril, is a draconic rule. It is destitute of any semblance of justice or humanity. It is cruel and wicked. It illustrates the subserviency of the American judiciary to the great corporations. . . . [It] puts the wealthy capitalist, corporate or unincorporate, upon the same equality in this respect, as that of the starving laborer, who must carry his meager dinner pail to his employment, no matter how dangerous it may be, in order to get a little food, clothing and shelter for his suffering family. . . . Those who can reconcile their consciences to the cold brutality of the general rule with reference to the servant accepting the risk, are at liberty to do so; I envy neither their heads nor their hearts.

ments of the employee’s prima facie case were not the principal legal issue. Instead, the legal battle raged over the famous trinity of defenses: contributory negligence of the employee, assumption of risk, and common employment. Yet more than bare theoretical interest attached to the articulation and justification of the appropriate legal rules, for the entire area of employers’ liability was widely thought to go to the heart not only of tort law, but also of the relationship between capital and labor in the production and exchange of goods and services. Indeed, by virtue of its apparent social and economic consequences, not to mention its sheer emotional importance, the problem could not in the end be contained within the judicial system. When workers’ compensation was introduced, it was done everywhere by legislation—legislation that in an explicit fashion adopted the views of the critics, not the defenders, of the nineteenth-century common law synthesis.

With the passage of the workers’ compensation statutes, the law of industrial accidents entered into a period of relative tranquility, as other problems of employment relations grew more pressing. The issues of collective bargaining, minimum wages, and employment discrimination surged to the fore. Workers’ compensation was relegated to a minor role, worthy of a problem that had been solved. As of late, however, workers’ compensation has returned to center stage; there have been not only strains within the system, but also sustained efforts to circumvent it. The exclusive remedy provision, which has long kept employers insulated from the rigors of the common law of tort, has come under mounting attack. The ebb and flow of the law of industrial accidents is itself an object of some interest, and the newfound concern with workers’ compensation invites a critical examination of its historical antecedents and intellectual foundations. It is to these topics that this Article is directed.

In order to make the organization clear, the Article discusses four different stages of development. The first concerns the common law rules of industrial accidents. The second concerns the early employers’ liability statutes for industrial accidents. The third concerns the adoption of the modern systems of workers’ compensation. And the fourth concerns the possible sources of their disintegration. The treatment here will not be complete on all matters of detail. Indeed, the internal operation of the modern workers’ compensation system is quite outside the scope of this Ar-
article. But I do hope that I can isolate and comment upon the central features of the history and of the debate, once thought closed, that the subject has generated.

I. Common Employment

The law of employers' liability, both in tort and in contract, is regarded as a child of the nineteenth century. In one sense this view is clearly correct. The first case reports on employers' tort liability date from 1837 in England with Priestley v. Fowler, and from 1841 in the United States with Murray v. South Carolina Railroad. In 1842 followed the landmark Massachusetts judgment of Lemuel Shaw in Farwell v. Boston & Worcester Railroad. Yet the employment relationship is itself of ancient origin. Its place in both England and America clearly precedes the industrialization of the last half of the eighteenth century, as well as the first reported workplace accident cases over a third of a century later. To say that there was no law on the subject before these epic cases were handed down would be an error. The utter dearth of cases upon the subject indicates, clearer than any judicial opinion could proclaim, an ironclad rule of breathtaking simplicity: no employee could ever recover from any employer for any workplace accident—period. In retrospect, the legal outcome could be delicately phrased by saying that the employer owed no duty of care to the employee, that the employee assumed the risks of all injuries associated with his employment, or that the employer was wholly immune, not unlike the sovereign himself, from legal accountability. Such conclusions were not stated, let alone justified. Silence said it all: an employee should be grateful for the opportunity for gainful employment. That he should receive any special legal protection on top of his good fortune was quite unthinkable. In a society in which disease and injury were rampant, and life itself fragile and short, the result should not come as too much of a surprise. Why

---

1 150 Eng. Rep. 1030 (Ex. 1837).
2 26 S.C.L. (1 McMul.) 385 (1841).
3 45 Mass. (4 Met.) 49 (1842).
4 It is unclear as an historical matter whether an employer could not be sued for a willful assault upon a worker. That cause of action could proceed wholly without reference to the employment relationship. Indeed, as the assault is deliberate, it removes the case from the category of industrial accidents, as is generally held today. See 2A A. Larson, Workmen's Compensation Law § 68.11 (1952).
should the legal system intervene on behalf of those fortunate enough to gain employment when there were countless others, far worse off, who would gladly trade places with them?

This utter dearth of precedent was a point well understood by both the English and American judges; indeed both Priestley and Farwell rely upon the absence of any precedent as a very strong reason why an employee could not maintain an action against an employer for the negligence of his fellow servant in the course of their common employment. Measured, therefore, solely by their outcomes, the two cases appear to underscore the uncompromising no-liability rule of the common law. No successful action for an industrial accident had been brought prior to these suits. Afterwards, nothing had changed.

It would, however, be a great mistake to measure the institutional importance of these decisions by the failure of the immediate plaintiffs to gain relief. In fact, the decisions in question ushered through the back door unheralded doctrinal changes that were in their own way as important as those wrought by the workers’ compensation statutes of a later day. The central point was the recognition—hinted at in Priestley, but carefully developed in Farwell—of the grounds on which an employee properly could sue an employer in negligence. The central distinction taken in Farwell was between the personal misconduct of a co-employee and the structural decisions that marked some defect in planning or organization of the employer’s business. Thus in Farwell, the employer escaped responsibility because the only negligence alleged was that of the employee’s fellow servant in failing to operate properly a switching device, itself a case of ordinary personal neglect. But the court intimated that a cause of action properly lay against the employer if there had been some defect in the design of the tracks, in the choice of equipment run on them, or in the selection and supervision of the employee. Applied in the corporate context the rule necessarily attached liability to at least some decisions of senior officials. Of all these commonplace sources of employer liabil-

---

7 150 Eng. Rep. at 1032.
8 45 Mass. (4 Met.) at 62.
9 Id.
10 In American courts this developed into a very formidable body of law about the vice-principal rule, a subject that received extensive judicial elaboration. See, e.g., 4 C. LABATT,
ity, there had been not a whisper in the decided cases. Yet now a
dearth of precedent, mysteriously, was no longer decisive. With
hindsight, it is fashionable to condemn both Priestley and Farwell
as vestiges of an ancient and uncaring system of employers' liabil-
ity. But, for better or for worse, they marked a genuine expansion
of employee's rights of action as against the employer. At an
institutional level, they were plaintiffs' decisions.

What is even more remarkable about Farwell is the level of so-
pphistication it brought to the basic problem. The decision, unlike
Shaw's unsatisfactory effort to rationalize negligence in Brown v.
Kendall, did not treat the rule of common employment in isolation.
To the contrary, Shaw made every effort to defend his result by
placing it in its larger legal structure. Consistent with this view,
he first drew the correct distinction between actions brought by
strangers and those brought by persons in consensual arrange-
ments with a defendant. The second class was a diverse mix, as it
contained actions brought not only by injured employees against

MASTER AND SERVANT 4141-665 (1913), which devotes over 500 pages to the subject. The
English development along these lines was curtailed by the important decision in Wilson v.
Merry, 1 L.R.S. & D. App. 326 (1868), which denied the employer's liability in the vice-
principal situation. The effect of the decision therefore was to extend the prohibition on
suits against the employer to situations in which supervisory personnel were involved. The
decision was widely castigated, and the reaction to it was in large measure responsible for
the passage of the Employers' Liability Act of 1880, 43 & 44 Vict., ch. 42. See Introduction
to T. Shearman & A. Redfield, A TREATISE ON THE LAW OF NEGLIGENCE (5th ed. 1898).

11 See, e.g., M. Horwitz, TRANSFORMATION OF AMERICAN LAW 210 (1977). Horwitz argues
that Farwell represented a retreat from the eighteenth-century view of contract whose "nor-
mative" elements allowed judges to determine not only the existence but also the fairness of
bargains. Horwitz also states that the adoption of the common-employment rule was "thor-
oughly determined by the intellectual impoverishment of counsel" who refused to argue the
point as a matter of tort law, conceding that the issues in question were purely contractual.
Id. Historically the point seems false, as the well-articulated opinion of Shaw suggests that
he followed his own considerable intellectual powers and would have reached the same conclu-
sion even if the point had been contested. The "normative" theory to which Horwitz
refers generally involved cases in which courts of equity were prepared to prevent the fore-
closure of mortgages or to set aside a sale of a widow's reversion because of grossly inade-
quate consideration. These matters are far removed from the employment context, where,
the employee seeks only damages, no basis for equitable jurisdiction can be found. For
an account of the eighteenth-century approach to contractual matters, see P. Atiyah, THE

12 Lawrence Friedman regards the opinion as "brilliant," see L. Friedman, A HISTORY
OF AMERICAN LAW 263, 414 (1973), but also as almost self-evidently wrong, see Friedman &
(1967).

13 60 Mass. (6 Cush.) 292 (1850).
their employers, but also by passengers against common carriers and guests against innkeepers, both common occurrences in an age innocent of suits against employers. As regards the stranger case, Farwell accepted the uncompromising position that vicarious liability held the employer responsible for the negligence of his servant within the scope of employment, a point of no little importance because it prevented railroads and other corporations from ignoring the full consequences of the conduct of individual employees. These rules were of enormous importance in all personal injury and property damage cases in the nineteenth century, and their importance remains undiminished today.

Yet there is no doubt that Shaw was correct as a matter of principle insofar as he claimed that actions as between strangers were governed by principles different from those applicable between parties to some consensual arrangement. For accidents arising out of consensual arrangements, Shaw took the position that the liability did not sound in tort, but rested, if at all, upon a promise by the defendant employer to compensate the plaintiff employee for the injuries sustained. In Farwell no express promise could be found; Shaw was thereby forced to address the murky question of the implied promise. This he did on the assumption that implied promises were of equal dignity with express ones. Yet major obstacles stand in the way of any coherent application of this view of the implied promise. There is no reason why express contractual provisions on loss allocation must be constant across different agreements of the same general class, much less across different classes of relations. The variation in perceived conditions is often met by a variation in contractual responses. In principle, implied contracts must mirror that complexity, but the possibilities of doing that are sharply limited within the ordinary context of litiga-

---

15 This basic insight explains why courts in their efforts to fashion a comprehensive code of products liability cannot settle upon a single standard for the entire myriad of cases that come within the domain. A strict liability rule, for example, is quite appropriate for construction defects, but wholly inappropriate for the unavoidable side effects of certain risky activities. The point is reflected in the Restatement (Second) of Torts § 402A comment k (1965), which makes as an exception to its general strict liability rule a warning requirement (based on negligence principles) for dangerous drugs. For a discussion of the general problem of a unified theory of liability in products cases, see R. Epstein, Modern Products Liability Law 147-51 (1980).
tion. To be sure, some judicial sensitivity to differences between different types of legal relationships is feasible. Yet it is far more difficult to show equal sensitivity to differences amongst individual cases within a given class of legal relations.

The point is well illustrated by the common employment relationships themselves. Both Priestley's butcher and Farwell's engineer brought actions as employees against their employers. Yet a priori there is no reason to expect that these two employment contracts, if express, would adopt the same rules of risk allocation. One business is very much larger than the other, and the ordinary principles of economies of scale may make it impossible for, say, the smaller firm to institute certain compensation programs that might be within the capabilities of a larger one. And the frequency of accidents, or the riskiness of the work, may be an important determinant in the level of investment in safety even in firms of the same size. If, therefore, one knows only that the plaintiff is an employee and the defendant an employer, there is insufficient information to identify the rule that properly allocates the risk of loss, and there is no reason a priori to assume that the same rule will work equally well in widely dissimilar contexts.

The first part of Shaw's analysis was, therefore, the least difficult, for in it Shaw had only to explain why passengers and guests were protected by a rule that held the carrier or innkeeper respectively responsible for all losses, save those that were vis major, or due to the act of enemy persons. Here he was aided because the possibility of any uniform rule was very great in both these contexts, for there is very little variation in the situations of both passengers and guests, at least in comparison with that of employees. One can, therefore, have some confidence that the proper implied rule will fit a very large percentage of the occasions it must govern. The real task lies in filling out the central contours of the general rule. In this case Shaw relied upon a sophisticated view of information costs. A rule of simple negligence would make it difficult, if not impossible, for the plaintiff to recover even for losses brought about by, for example, an innkeeper's treachery—as where he acted in cahoots with a band of robbers. The imposition of the higher standard of care in situations where a plaintiff, be he a passenger or guest, was so ignorant of the possible perils that he could do

---

16 See text accompanying notes 86-87 infra.
little for his own welfare instilled incentives for caution in the defendant. At the same time a limited escape hatch for visible and extreme hazards—vis major, enemy acts—spared defendants from liability for self-authenticating losses that (i) they did not cause, (ii) they could not be expected to prevent, and (iii) might well have occurred even if the plaintiff had not trusted himself into the care of the defendant carrier or innkeeper.

There are few, then or now, who are disposed to criticize Shaw's treatment of either the innkeeper or the carrier. But for our purpose, its very soundness, coupled with the extensive system of liability that it entailed, exposes his treatment of employers' liability to criticism. Shaw, it must be recalled, did not take the easy course and say that the liability rules for carriers and innkeepers could be carried over without difficulty to employers. To the contrary, he argued that the justifications for extensive liability in those cases demanded a far more restrictive regime of employers' liability. As before, as a first approximation, liability should fall on the party with the best knowledge and ability to prevent the loss in question. In the workplace, however, Shaw insisted that the proper person was often the employee. The passenger and the guest may have entrusted their safety to the care of others, but the employee knows that for his own self-preservation he must care for himself, for which he is equipped with knowledge of the circumstances in the field and with the ability, alone and with peers, to exert moral suasion and social pressure upon his fellow servants. These arguments have a strong bite in Priestley. The two servants were face to face, and the employer was nowhere to be found. For the common precautions needed there is little that an employer can do either to instruct workers (of what they already know) or supervise them (when they are on their own). The forces of self-preservation may have been greater guardians of safety than any liability rule. Nor should the incentive upon the employer be forgotten: while he was legally off the risk, if the accident level in the firm became too high, further wage increases would be needed to attract and hold the best workers.\(^\text{17}\) In addition, an employer (especially a railroad)

\(^{17}\) On this point, I think it is possible to disagree with Professor Rabin when he argues against the fellow servant rule on the ground that it imposes no incentives upon the employer to provide a safe workplace. See Rabin, *The Historical Development of the Fault Principle: A Reinterpretation*, 15 GA. L. Rev. 925, 940-41 (1981). Rabin's position is that
would have some incentive to take care of his employees if only as a necessary by-product of the discharge of his obligation to care for passengers and strangers.

There is still, however, the question of the variation within the class of employment contracts. This was presented in Farwell, as Shaw had to decide whether the fellow servant doctrine reached cases in which the two servants were employed in "different departments"—the point pressed by plaintiff as a reason to limit, but not abandon, Priestley. Yet Shaw effortlessly finessed the point on two separate grounds. First, he noted that any effort to limit the fellow servant doctrine to employees of the same department required an independent account of what constituted a department. For him the administrative costs of policing the distinction were far greater than any perceived benefit that might be derived from its use. That consideration could also be used to argue that, if the common-employment doctrine was wholly inappropriate for employees of different departments, then it should be abandoned entirely. Yet this Shaw refused to do, in part because he feared the administrative complications that suits against employers would entail, and in part because he thought that any additional risks were offset by higher wages—pointing out that Farwell had received higher wages when he moved from his former job as a mechanic to the more dangerous employment as an engineer.

---

1 45 Mass. (4 Met.) at 52. The twists and turns of the "different department rule" are impossible to summarize within a short space. It should be sufficient to note that Labatt devotes about 70 pages to the issue. See 4 C. Labatt, supra note 10, at 4071-140. In one place he writes that "while it is agreed on all hands, that, under some circumstances, a difference in department is a ground allowing recovery, there is a radical and irreconcilable divergence of opinions as to what the circumstances are that will, in this point of view, exclude the defense of common employment." Id. at 4102. Posner's treatment of the issue is a bit more cavalier: "Under the 'different department' rule, the employer was liable for the negligence of an employee in a different department of the company from the one in which the injured man worked." Posner, A Theory of Negligence, 1 J. Legal Stud. 29, 67 (1972). No cases are cited, and his conclusion is inconsistent with Farwell, which remained heavily influential throughout the nineteenth century.

18 45 Mass. (4 Met.) at 59. For a modern endorsement, see Posner, supra note 18, at 70. Posner also accounts for the assumption of risk argument on the ground that most workers are "risk preferring," id. at 71, which goes against the common assumption that most individuals are risk neutral or perhaps even risk averse. In any event the divergence of risk
Yet here too the argument moves a bit too fast for its own good: the wage differential may persist, but its origins are unclear. It could be for the increased hazards of the work, but Shaw made no effort to show that these were attributable to the conduct of fellow servants and not to other sources, including, for example, employees of other railroads that used the common tracks. But without this showing, the premium paid need not be tied to the employee's waiver of the right to sue the employer for harm caused by the fellow servant. Moreover, if the fellow servant presented an increased risk to the plaintiff, to treat the wage differential as a premium presupposes that the plaintiff's damage action against a fellow servant was not well conceived, an issue not resolved until the case at hand. In any event some differential could well exist even if the action were allowed on the grounds that, with hazards of litigation and calculation of damage awards, it would not make the plaintiff whole. Still other explanations for the wage differential are possible—it could be that the plaintiff moved to a more skilled and valued job. Any inference from the wage differential to the selection of the proper liability rule becomes somewhat attenuated.

The nub of the criticism of Farwell is that a gap-filling rule created by a court should not be treated as having the same force or commanding the same respect as an express contract. Yet the point was often missed in the cases that followed. Thus in Hutchinson v. York, Newcastle & Berwick Railway, the first English case to apply the doctrine of common employment to railroads, the distinction between harms to strangers and harms to fellow servants was again taken; and again it was asserted that contract principles precluded the action by the servant against the employer. But the noncontractual elements of the rule show through the decision of Baron Alderson; the servant "knew when he was engaged in the service, that he was exposed to the risk of injury, not only from his own want of skill or care, but also from the want of it on the part of his fellow-servant; and he must be supposed to have contracted on the terms that, as between himself and his preferences among employers and employees makes it difficult to believe that a single rule is ideal from a contractual point of view. For criticism of the Posner position, see Rabin, supra note 17, at 940-41.

master, he would run this risk." The "must be supposed to," however, says it all. Here there was no evidence that the terms of the agreement were stated by the employer to the employee as a condition of his employment. And there is no principled explanation of why the higher wage without the liability in question is necessarily preferable to a lower wage with certain forms of protection. It is in my view a major oversimplification to urge, as does Professor Posner, that the "courts in these cases are simply enforcing the employment contract" or that the common law rules vindicated the position of the risk-prefering workers. The rules of industrial accidents may have been developed by judges who believed in freedom of contract. But the rules themselves were not contractual, at least as explicit contracts are contractual.

The obvious equivocation between implied in fact and implied in law has always made the rule of common employment an obvious object of attack. Thus Prosser, always the touchstone of orthodox wisdom, has written:

The cornerstone of the common law edifice was the economic theory that there was complete mobility of labor, that the supply of work was unlimited, and that the workman was an entirely free agent, under no compulsion to enter into the employment. . . . The economic compulsion which left him no choice except starvation, or equally dangerous employment elsewhere, was entirely disregarded.

It should be evident that these criticisms go far beyond the proper treatment of the implied term and reach ultimate issues of public policy. Yet while these criticisms are more fundamental, they are also less persuasive. In particular, there is no need to make a set of narrow assumptions about market structure and employee behavior in order to defend a rule of contractual freedom in employment contexts. Even on utilitarian grounds, it is only necessary to show that it yields better results than the next best set of

---

21 Id. at 154.
23 Posner, supra note 18, at 71.
contractual implications. It would, of course, be a happy set of circumstances if knowledge of employees were perfect and the mobility of labor were wholly unimpeded. We all know the many virtues of perfect competition in a world without transaction costs. But the freedom of contract rule is not defensible only under such restrictive assumptions. A more complicated set of market conditions within employment markets makes any legal rule more difficult to defend, including one that presumptively permits actions for negligence against employers. So long as there is some mobility of labor and some knowledge of employees, it may well be that a rule that presumptively protects the employer from a suit is better than one that exposes him to it as a matter of course. To take Prosser seriously, one would wonder why employers did not pay zero wages and why they ever made any investments in safety at all. To be sure, there was some sharp economic compulsion, but this is hardly removed by driving employers from the marketplace. The "author" of the compulsion was the hard conditions of life and the low standards of living from which the technological advances of the nineteenth century promised the first real hope of escape. Was it really preferable to remain on the farm, to eke out a hostile living from the soil, and to have no employer to sue in the event of injury or death?

With this said, the defense of the rule of common employment takes a more limited form. The key point is that the enormous variation between employments makes it quite impossible for any court to develop a set of rules that will withstand the pressures of prolonged litigation. The ceaseless elaboration of the tort law of industrial accidents shows how fruitless it is to think that any one set of rules will do suitable service in the many separate conditions to which it will have to be applied. Yet if no ideal surrogates for contract can be developed, it is vital that the judicial rules be amenable to modification by private agreement. The rule of common employment with its commitment to freedom of contract allowed such modification. While transaction costs in some employment contexts may be high, the prospects in question are not wholly idle, as our discussion of the English Employers' Liability Act of 1880 indicates.

---

See Posner, supra note 18, at 75.
II. Employers' Liability Acts

The general dissatisfaction with the rule of common employment, especially in the uncompromising form it assumed in Wilson v. Merry, led to its legislative repeal first in England and then in the United States. The repeal of common employment did not, however, usher in the modern system of workers' compensation, for the initial response was the adoption of an intermediate position that imposed upon the employer a qualified form of negligence liability. The first statute of this sort was the English Employers' Liability Act of 1880, which served as the model for statutes in a good number of American jurisdictions.

The central features of the 1880 statute can be summarized as follows: section 1 of the statute enumerated five types of circum-

---

46 See note 10 supra.
47 43 & 44 Vict., ch. 42.
49 In full, the first two sections of the Act provide:
1. Where after the commencement of this Act personal injury is caused to a workman
   (1) By reason of any defect in the condition of the ways, works, machinery, or plant connected with or used in the business of the employer; or
   (2) By reason of the negligence of any person in the service of the employer who has any superintendence entrusted to him whilst in the exercise of such superintendence; or
   (3) By reason of negligence of any person in the service of the employer to whose orders or directions the workman at the time of injury was bound to conform, and did conform, where such injury resulted from his having so conformed; or
   (4) By reason of the act or omission of any person in the service of the employer done or made in obedience to the rules or byelaws of the employer, or in obedience to particular instructions given by any person delegated with the authority of the employer in that behalf; or
   (5) By reason of the negligence of any person in the service of the employer who has the charge or control of any signal, points, locomotive engine, or train upon a railway,

   the workman, or in case the injury results in death, the legal personal representatives of the workman, and any persons entitled in case of death, shall have the same right of compensation and remedies against the employer as if the workman had not been a workman nor in the service of the employer, nor engaged in his work.
2. A workman shall not be entitled under this Act to any right of compensation or
stances in which the traditional common law rules of liability were displaced. The first of these dealt with defects in the condition of ways, works, machinery, or plant; the next three dealt with cases in which the employee acts under the supervision or direction of other workers, thereby reversing the rule of Wilson v. Merry; and the last was directed to situations like that in Farwell, where one worker is injured by the "negligence of any person . . . who has the charge or control of any signal points, locomotive engine, or train upon a railway." With respect to these five categories of cases, but to none other, the statute provided in delphic form that the workman in case of injury, or his representative in case of death, "shall have the same right of compensation and remedies against the employer as if the workman had not been a workman of nor in the service of the employer, nor engaged in his work."

The second section of the Act then specified a set of exceptions to the basic rules of liability contained in the first provision. The first exception demanded that any defect in condition set out in section 1 had to be attributable to the negligence of the employer or his servant. This clause need not have been drafted as an exception as it was possible to incorporate the negligence requirement into the principal section, as was done in subsections 1(2), 1(3), and 1(5). The second exception was in fact a proviso to section 1(4)

remedy against the employer in any of the following cases; that is to say,

(1) Under sub-section one of section one, unless the defect therein mentioned arose from, or had not been discovered or remedied owing to the negligence of the employer, or of some person in the service of the employer, and entrusted by him with the duty of seeing that the ways, works, machinery, or plant were in proper condition.

(2) Under sub-section four of section one, unless the injury resulted from some impropriety or defect in the rules, byelaws, or instructions therein mentioned; provided that where a rule or byelaw has been approved or has been accepted as a proper rule or byelaw of one of Her Majesty's Principal Secretaries of State, or by the Board of Trade or any other department of the Government, under or by virtue of any Act of Parliament, it shall not be deemed for the purposes of this Act to be an improper or defective rule or byelaw.

(3) In any case where the workman knew of the defect or negligence which caused his injury, and failed within a reasonable time to give, or cause to be given, information thereof to the employer or some person superior to himself in the service of the employer, unless he was aware that the employer or such superior already knew of the said defect or negligence.

43 & 44 Vict., ch. 42.

50 Id. § 1(5).

51 Id. § 1.
that stated that no rule or bylaw should be regarded as defective if approved by the Parliament or its agents. The last of the exceptions, and the one that caused the greatest difficulties in the construction, went to the issue of assumption of risk, as inferred from the failure of the employee to give notice to the employer of defects of which he had knowledge and the employer did not.

Thereafter the statute set out the limited amounts of money that could be recovered—three years' wages were a maximum; the periods of time in which the actions could be brought; the mechanics of payment and collection; and provisions governing the trial of suits and the notice of claim and the like.

In dealing with the statutory scheme, this Article addresses one major theme that goes to the status of employers' liability under the statute: was it possible for the employer and employee to contract out of the statute by express agreement? This question was raised shortly after the passage of the statute in Griffiths v. Earl of Dudley. There the decedent fell to his death in a pit accident because of a defect in the equipment used to take miners in and out of the pit shaft. The defect in the condition of the plant fell squarely within section 1 of the statute, and it was attributable to the negligence of an inspector of machinery also in the defendant's employ. The sole defense raised against the recovery of damages was that the employee had agreed as a condition of employment to waive all the rights accorded him under the Employers' Liability Act. The full terms of the agreement must be outlined, for at issue in the case was not a simple provision that purported merely to restore the former judge-made rules of common employment; rather, at issue was a detailed arrangement that created its own compensation scheme in the event of employee injury or death.

The agreement first noted that there existed prior to the passage

Limit of sum recoverable as compensation.

3. The amount of compensation recoverable under this Act shall not exceed such sum as may be found to be equivalent to the estimated earnings, during the three years preceding the injury, of a person in the same grade employed during those years in the like employment and in the district in which the workman is employed at the time of injury.

Id. § 3.

Id. § 4.

Id. § 5.

Id. §§ 6-7.

9 Q.B.D. 357 (1882).
of the Employers’ Liability Act a benefit society that operated a “field box.” Contributions to the field box were made by levying weekly assessments against the individual workers as a condition of their employment; these were then matched by an equal contribution from the employer, and payments from the box were to be made by the employer in his sound discretion for any worker who was injured or killed “in the course of his employment.” The details of the operation of the scheme were not revealed in the report, but there is no evidence that recovery from it depended upon proof of the negligence of the employer or that it could be barred by evidence of ordinary misconduct, be it by the injured party or by his fellow servants. In sum, the structure of the system looks very much more like that of modern workers’ compensation statutes than either the restrictive common law liability rules in place before the passage of the Employers’ Liability Act or the negligence system that the Act itself instituted.

Following passage of the statute, the employer posted notice that he expected that the old arrangement for compensation would continue and that in order for that to be the case, all benefits under the Act had to be waived as a condition for employment. The notices so posted were read by the decedent, who in turn made his regular contributions into the box, so that an agreement by notice and conduct can be easily inferred under the ordinary rules of contract formation. The question before the court was whether this arrangement was barred by the statute, which had no words to that effect, or whether the statute could be displaced by the agreement, now express in its formation and clear in its terms.

The decision in the case, which reflected a strong and explicit bias toward freedom of contract, was that opting out was not only possible, but entirely proper. “There was no suggestion that the contract was induced by fraud, or by force, or made under duress, and it was not a naked bargain made without consideration, for the defendant contributed an amount to the club equal to the whole amount of contributions from the workmen.” Again: “It is said that the intention of the legislature to protect workmen against imprudent bargains will be frustrated if contracts like this one are allowed to stand. I should say that workmen as a rule were per-

37 Id. at 358.
38 Id. at 362.
fectly competent to make reasonable bargains for themselves.\textsuperscript{39} And last, "I think great injustice would result [if the contracts were not respected], because the workman might obtain the benefit of the contract for years in the form of higher wages to cover the risk of injury, and then claim full additional compensation when he was injured."\textsuperscript{40} I am not ashamed to say that I agree.

Moreover, Griffiths is not the only case of this sort in the reports. Clements v. London & Northwestern Railway\textsuperscript{41} presented for judicial consideration a voluntary compensation scheme that reads like a blueprint of the modern workers' compensation rules. The negligence of the employer was immaterial to the question so long as the injury was work-related; and only the employee's rightful neglect or gross negligence allowed, but did not require, the society to eliminate the benefits in question. In addition, disputes between the two were to be resolved not by judicial action, but by arbitration between the parties.

A second feature of Clements was that the action was brought by an infant in the service of the railroad, for whom the usual rule allowed enforcement of the contract only if its provisions on balance were fair in their essential features.\textsuperscript{42} Here a unanimous court of appeals—one not indisposed to protect infants—upheld the agreement under attack. The arguments used by Lord Esher—itself a point of no small significance—sound like a brief

\textsuperscript{39} Id. at 363.
\textsuperscript{40} Id. at 364.
\textsuperscript{41} [1894] 2 Q.B. 482.
\textsuperscript{42} Id. at 495.

There were other statutes of more limited scope that moved in the same direction, which I cannot consider here, but which doubtless play an important part in the general evaluation of the law of industrial accidents. See, e.g., Act of Mar. 5, 1856, no. 103, 1856 Ga. Laws 154 (repealing common employment defense as against railroad employees and allowing them to recover if free of negligence); see also Act of Apr. 8, 1862, ch. 169, § 7, 1862 Iowa Acts 196 (to the same effect). Note that these systems that introduce negligence and contributory negligence are, like the Employers' Liability Act, a far cry from the workers' compensation statutes and are in any event not necessarily (or even clearly) superior to the original common law rules.

\textsuperscript{43} With Clements contrast Flower v. London & Nw. Ry., [1894] 2 Q.B. 65, where the court refused to enforce a contract against an infant that exonerated the employer from all liability in exchange for free transport to work.

\textsuperscript{44} The significance stems from Lord Esher's role in the judicial battles over the proper construction of the Employers' Liability Act. In opposition to Bowen, L.J., he always urged a broad interpretation of the Act, including a narrow version of assumption of risk. Lord Esher urged this interpretation in his powerful dissent in Thomas v. Quartermaine, 18
for the legislative adoption of the rule in question:

If there were no such contract, he could not obtain compensation, unless by agreement with his employers, without bringing an action either in the superior court or the county court, and in that action he would be exposed to the risk of being unable to prove that the accident was the result of negligence of some one for whom the company were responsible. The injuries might, for instance, have arisen from concealed defect of machinery not known to the company, or by pure accident not brought about by any negligence on the part of the company's servants. The burden of proving that it was otherwise would have been on the plaintiff, and that is a burden which often cannot be supported. Even if the plaintiff were successful in shewing this and obtained judgment, and the defendants had to pay his costs, it is a matter of common knowledge that the plaintiff would have to incur extra costs beyond those he would recover. Such extra costs would have to be paid out of the damages which he would recover; and we all know that in a majority of the cases in which only small damages are recovered those damages are seriously encroached on in meeting the extra costs.

The risk of non-success owing to difficulty of proof, and the risk of obtaining but small advantage from a successful action, are both obviated by this agreement, under which, even if it is clear that there is no legal claim which could be enforced against the company, he is still entitled to compensation.\(^4\)

The bargain in question was not only expressly made, but, given what the employee obtained as well as what he surrendered, was fair in all its circumstances. That result, it might be added, could have been reached by another route, the nondiscrimination principle. The infant was admitted to the association on the same terms as adult members, so that he received the protection afforded all those competent to bargain for themselves. As in Griffiths a volun-

---

Q.B.D. 685 (1887), and his majority opinion in Yarmouth v. France, 19 Q.B.D. 647 (1887). He also took a very broad view of what counted as a product defect. See Walsh v. Whiteley, 21 Q.B.D. 371 (1888), discussed at text accompanying notes 53-57 infra. The central point is that Esher did not question the validity of the explicit contractual arrangements although he was hostile to the assumption of risk defense.

\(^4\) Clements, [1894] 2 Q.B. at 489-90; see also note 57 infra.
tary organization rejected both the common-employment rule of
the common law and the statutory negligence rules designed to re-
place it. At some considerable cost the private parties instituted
instead a full-fledged antecedent to the workers’ compensation
rules.

The pattern revealed in both Griffiths and Clements is impor-
tant for what it says about the ability of the courts to determine
the efficient provisions in the absence of express agreements. Pro-
fessor Posner, in his examination of industrial accidents, properly
observed that some limited evidence of the efficiency of common
law rules can be found in the failure of parties to contract out from
them. As regards the rules of assumption of risk and common em-
ployment, he found no evidence in the cases of contracting out in
the American context. The English experience suggests quite oth-
erwise. Yet here it is important to have some sense of the fre-
quency and distribution of the voluntary movement to workers’
compensation. The plans themselves come from two indus-
tries—railroads and mining—where the rate of industrial accidents
relative, say, to retailing, would be quite high. In both cases the
plans had a finished, comprehensive look that suggests that they
were not some isolated peculiarity but the product of a gradual
and perhaps widespread trend. A recent unpublished paper by
Veljanowski of the London School of Economics provides some in-
dication that the plans themselves had, at least in the collieries,
broad adoption. He notes that something under twenty-five per-
cent of the mines had such plans in force and that these were con-
centrated in Lancashire, where mining was the most dangerous and
the need for some positive response to the general question of acci-
dents was most pressing. One possible way to interpret his data is
as evidence that the compensation programs were generally supe-

44 Posner, supra note 18, at 74. In his Article, however, he notes, but without references:
“Evidently it was the practice of large employers, such as railroads, to provide medical
treatment to injured employees free of charge and to compensate them, during a limited
period of disability, for lost wages; in exchange the employee would agree not to bring a tort
action.” Id. at 83. If the agreements in question were entered into before the accident, then
they look like the adoption, perhaps informal, of a workers’ compensation system; if after
the accident, they look like the settlement of a tort case. The differences are not explored by
Posner.

manuscript) (available in University of Chicago Law School Library).

48 Id.
rior to the common law rules, but that the costs of their introduction were such that they were adopted only in those places where they were most needed. In this regard the great contribution of the English common law judges in response to the Employers' Liability Act was not to erect obstacles that limited contractual freedom.\(^6\) Both plans presented to the courts provided the worker with a quid pro quo for the release of his rights under the Act that was not simply buried into the wage as part of an ineffable premium. Whether the courts would have gone the full extreme and allowed the contract defense without the substitute program must remain something of an open question.\(^5\) What can be said with confidence is only that in many circumstances the common law rules were inferior to the contractual alternatives.

Why should this be so? Here it is instructive to look at the points raised by Lord Esher in *Clements*, for they represent a critique of the Employers' Liability Act that in large measure applied with equal force to the common law rules that his points had modified. The original coverage provisions contained in section 1 of the Act may have appeared at first glance to sweep within the statute virtually all industrial accidents. But actually, in what in retrospect seem odd, if not indefensible, decisions, the separate heads involved were treated as *limitations* upon the rights of employees that could be triggered under the Act. A few illustrations make the basic point. First, in *McGiffin v. Palmer's Shipbuilding & Iron Co.*,\(^5\) the decedent was killed when a car containing a heated ball overturned and fell upon him while he was pushing it along a track. The track was obstructed by a piece of material used to line a furnace. The claim for damages under the Employers' Liability Act was denied on grounds that a "defect in the condition of the way" was not so broad an expression as to cover cases of obstruction, but reached only those in which the tracks themselves were of

\(^6\) Note that the American statutes on the subject did have express prohibitions on contracting out. *See, e.g.*, Federal Employers' Liability Act, ch. 149, § 5, 35 Stat. 65, 66 (1908); Federal Employers' Liability Act, ch. 3073, § 3, 34 Stat. 232 (1906); Act of May 14, 1887, ch. 270, § 4, 1887 Mass. Acts 899.

\(^5\) The quid pro quo has loomed large in the American constitutional treatment of workers' compensation. *See, e.g.*, New York Cent. R.R. v. White, 243 U.S. 188 (1917), which in part upheld the statute because of the quid pro quo (in the form of the exclusive remedy provision) received by the employer.

\(^5\) 10 Q.B.D. 5 (1882).
defective manufacture or repair, and perhaps others in which they were rendered slippery with ice. It is difficult to think of a principled reason why the Act should be extended to one class of defects while withheld from another. Even if it is objected that the innocent words of a statute were construed in unintended ways that denied the recovery, it still remains true that the defect and negligence language functioned as words of limitation as well as those of entitlement. Marginal adjustments in their meaning could serve to defeat ordinary expectations in ordinary cases.

*Walsh v. Whiteley* illustrates the same point. The plaintiff’s job required him to place a band upon the moving wheel of a carding machine. The wheel itself contained holes in order to render it lighter and therefore susceptible to easier motion. While operating the machine, the plaintiff caught his thumb in one of the holes. Recovery was denied—Lord Esher dissenting—after an exhaustive discussion of the nature of and relationship between danger, defect, and negligence, with the majority taking the view that if the best possible machines were regarded as “defective,” the term would be drained of all content in the law. Lord Esher’s dissent stressed that the dangerous condition of the machine necessarily made it defective. The discussion itself is not dissimilar from that taking place today in connection with the modern law of products liability, where the treatment of the issue, with the passing of the open and obvious defect test, is as troublesome as it was nearly one hundred years ago. But the case neatly illustrates the perils of negligence litigation, even where there is no systematic bias in the administration of the Act or ineptitude by the judges charged with its construction.

A similar message comes up, moreover, when one deals with the

---

52 Id. at 8-10.
54 Id. at 378-80.
55 Id. at 375-76.
56 For the modern rejection of that defense, see Micallef v. Miehle Co., 39 N.Y.2d 376, 348 N.E.2d 571, 384 N.Y.S.2d 115 (1976), which I have criticized, see R. Epstein, *supra* note 15, at 144-45.
57 The same analysis, again at the constitutional level, was conducted in connection with the Supreme Court’s analysis of the Price-Anderson Act. The quick and reliable remedy provided by the statute was held preferable to the uncertain rights of claimants at common law, even with the maximum cap on damages that the statute provided. See Duke Power Co. v. Carolina Environmental Study Group, Inc., 438 U.S. 59, 88 (1978).
treatment of defenses under the statute. Here it is quite clear that the draftsman himself was notoriously opaque in his specification of the consequences of a breach within the first section. Its specification is wholly negative: the workman is not to be treated as a workman but as someone else. The question then is: Whom? One possibility is that the workman could be treated as a stranger, but this would be odd because typically he is injured on the premises of the employer. Another possibility is to say that stripped of his identity as a workman he could be regarded as a trespasser, who would in consequence receive, if anything, less protection than he did before the passage of the Act. In *Thomas v. Quartermaine*,* 5* Bowen, L.J., in a powerful opinion, virtually read the defense out of the Act by allowing the employer a directed verdict whenever the worker continued on the job after he had knowledge of the condition that caused the injury. In my view the construction of the statute offered by Lord Esher was on balance preferable, as section 2(3) deals expressly with the matter,* 60* rendering it unnecessary to draw implications about the proper treatment of the issue from the other portions of the statute. The very next year, in *Yarmouth v. France,* 61 he was able to do an end run around *Thomas v. Quartermaine,* and his views were eventually adopted by the House of Lords in *Smith v. Baker & Sons.* Yet even here it must be remembered that Esher’s position did not bar assumption of risk in cases brought under the Employers’ Liability Act, but only left the matter to the discretion of the jury. It is another instance in which general legislation was unable to elim-

---

* 5 18 Q.B.D. 685 (1887).
* 60 This is also the opinion expressed by Francis Bohlen in his classical treatment of assumption of risk. See Bohlen, *Voluntary Assumption of Risk,* 20 Harv. L. Rev. 91, 100 (1906).
* 61 19 Q.B.D. 647 (1888), which, as Bohlen notes, was not followed in America. Bohlen, *supra* note 60, at 97 n.1.
* 62 1891 A.C. 325.
inate the great degree of factual uncertainty in individual accident cases. The prolonged and heated debate over assumption of risk affords yet another explanation of why individual workers were well advised to abandon their rights under the Employers' Liability Act in order to secure the relatively fixed and certain benefits that the alternative voluntary plans afforded. While the common law rule of common employment could be subject to powerful criticism, the same could be said of the Employers' Liability Act that replaced it.

III. WORKMEN'S COMPENSATION ACT OF 1897

A. The Act

The next stage in our history is concerned with the introduction of the modern workers' compensation statutes. Here too the developments in England were in advance of those in America, and the language of its statute—the Workmen's Compensation Act of 1897— served as the model for the subsequent American statutes. To understand the strengths and weaknesses of this statute, it is important to view it not in isolation, but in comparison to the earlier Employers' Liability Act, which it supplanted. Here several points of contrast between the two enactments merit special attention. The first of these goes to the basic coverage formula; the second to the status of contracting out under the statute; and the third to an issue that could not have arisen before the adoption of workers' compensation—the preservation of tort actions, whether against the employer or against some third party not covered by the Act.

On the matter of coverage it will be recalled that the Employers' Liability Act did two things: first, it kept a general negligence standard, and second, it identified certain particular classes of injuries covered by the statute. The newer statute took a very different approach to the basic coverage question. Negligence was removed as a condition of liability; and the particular types of injuries were replaced with a single comprehensive formula still in use today that refers only to "personal injury by accident arising out of and in the course of employment." As drafted, the statute did not ap-

---

*60 & 61 Vict., ch. 37.
*Id. § 1(1).
ply to all forms of employment, but only to work "in or about a railway, factory, mine, quarry or engineering work, and to [con-
struction of buildings]," leaving, it appears, ordinary retail per-
sonnel and household servants, for example, outside its purview.
The various defenses that might have been available at common
law or under the Employers’ Liability Act were also swept away,
for once the injury came within the statute, the employer was lia-
ble, except for a mandatory disallowance in cases in which it was
proved that the injury to a workman was attributable to "the seri-
ous and wilful misconduct of that workman." For the trades and
occupations covered, we have in essence the modern system of
workers’ compensation.

The second feature of especial importance in the Act was its
treatment of contracting out. The gap in the Employers’ Liability
Act was filled by a detailed scheme that controlled every aspect of
the process. Contracting out was allowed, in principle, but heavily
regulated in fact. One condition was that the contract in question
must be "on the whole not less favorable to the general body of the
workmen and their dependents than the provisions of this Act." For another, the scheme in question could only be put in place
after certification by the "Registrar of Friendly Societies," after
opportunity to hear from both employer and workmen. The certi-
fication in question was for a minimum five-year period, but it was
subject to revocation if, upon examination of complaints, the
plan was found in its administration to have fallen below the re-
quired standards and no steps were taken within a reasonable pe-
riod of time to correct the defects. In addition, it was no longer
possible to require "workmen to join the scheme as a condition of
their hiring."

Last, the statute did not eliminate the common law of tort
against the employer. "When the injury was caused by the per-
sonal negligence or wilful act of the employer, or of some person
for whose act or default the employer is responsible," the worker

---

**Id. § 7(1).**

**Id. § 1(2)(c).**

**Id. § 3(1).**

**Id.**

**Id. § 3(2).**

**Id. § 3(4).**

**Id. § 3(3).**
had the option against the employer to "take the same proceedings as were open to him before the commencement of this Act"—which presumably covers both common law actions and suits under the Employers' Liability Act—even though he could not thereby obtain a dual recovery. That option was made enormously attractive to the employee. If he chose his tort suit against the employer and lost, he could still, if the facts disclosed so warranted, receive his compensation benefits, less only the employer's costs of defense in the original tort action. The action against the third party was on somewhat less attractive terms, as the worker was forced to an election between the tort action against the third party and his compensation claim against the employer. Matters were made more complex as the employer himself, if required to pay compensation under the Act, could receive indemnification against the third party, although not it appears for the full extent of the tort damages.

As drafted, the 1897 statute answered the question left open under the Employers' Liability Act. The quid pro quo given the worker for a release from the statute could not be a simple increase in wages: some substitute disability benefit was required. Yet here the remaining provisions of the statute frustrated the voluntary plans. Note for example that the plans in both Griffiths and Clemments had to fail because both required all employees to contribute to the plan as a condition of employment. Prohibiting the employer from insisting upon uniformity was a crippling blow to the voluntary plans. The risk of adverse selection by employees against the firm was greatly increased, as individual workers were able to trade on their personal knowledge in deciding whether to opt in or out. Administrative costs were likewise pushed upward, as the employer had to absorb the costs of running two separate plans. All the while, the employer ran the risk that its plan would be struck down after complaint, leaving utterly undetermined the status of the contributions previously made to the plan and the payouts.

\[72\] Id. \S 1(2)(b).
\[75\] Id. \S 6.
\[71\] Id. Suits of this sort were relatively infrequent given the restrictive view of products liability law. See text accompanying notes 88-91 infra. And the larger question of coordination can be solved in a wide variety of other ways today, since the rules of comparative negligence permit many separate parties to share in the losses under, of course, widely expanded liability rules. See note 98 infra.
under it—a point that weighed heavily on Field, J., in *Griffiths*.

Difficulties were further compounded as the plans were not self-executing, but required the employer to obtain a prior approval at a hearing at which all opponents to the plan could protest. No voluntary organization could thrive under a set of external regulations wholly inconsistent with its informal modus operandi. In a very short time, the statutory protections overwhelmed the voluntary norm. What little litigation occurs under the 1897 statute (and its successors) is directed to the liquidation of the voluntary plans.

**B. The Compensation Bargain**

How then could the workers' compensation programs succeed? The answer requires us to address the important question of what consequences follow once a given set of injuries—here industrial accidents—is effectively removed from the market. The regulation is a barrier against the voluntary ideal. But its costs cannot be determined solely from knowledge of its existence. In addition, we must know the size of the deviations from the voluntary norm demanded by the compulsory scheme. The earlier history of the voluntary plans under the English Employers' Liability Act is evidence that in some critical situations workers' compensation better approximated the consensual norm than the common law rules of negligence, with their complex baggage of contributory negligence, fellow servant, and assumption of risk. At the structural level, the costs of the mandatory imposition may well have been low, for it only adopted a pattern already in widespread use.

In this connection the basic structure of the bargain was well understood with the passage of the Act. The broad coverage formula eliminated the need to determine negligence on both sides and assumption of risk—all inquiries with a high degree of uncertainty. In exchange for the broad coverage formula, the workman received a level of compensation that, by design, left him worse off than if the injury itself had never taken place. The low levels of the benefits doubtless proved nettlesome to workers after injuries. But to concentrate on that point is to miss the central role. First, low damages help keep down the overall costs of the plan, which will induce employers to continue to hire labor. Second, low bene-

---

75 See text accompanying note 40 *supra*.

76 See, e.g., Howarth v. Andrew Knowles & Sons, [1913] 3 K.B. 675.
fits help prevent fraud against the plan, as there is less to gain by pretending that an injury, or its consequences, is work-related. Third, the low awards create additional incentives upon the worker for self-protection and therefore act as an implicit substitute for assumption of risk and contributory negligence.

The question of incentives deserves a fuller discussion. In particular, it is instructive to compare the incentives created by the compensation system with those that are found under the negligence regime in place around 1900—a system in which negligence was required to state a cause of action and contributory negligence was treated as a complete bar. One defense of the negligence system on efficiency grounds, elaborated by Landes and Posner, is that it places the full costs of an accident on both parties to the transaction, here assumed to be both rational and self-interested actors. As I understand the argument, it looks first to the position of the defendant and then to that of the plaintiff. Within the workplace the employer will take due care because he knows that the employee (in order to recover damages) will also take good care, lest his action be barred. If, therefore, the employer does not, he will be forced to bear the full loss, for which, by definition, cost effective precautions are a cheaper substitute. Now once the employee knows that the employer will take care, he too will take care: by definition his costs of avoidance are less than the full costs of injury that would otherwise be visited upon him. In one sense the theory is very odd, for as a first approximation it predicts that while there will be accidents (i.e., those not worth avoiding), there will not be negligence. As that conclusion is falsified by the most commonplace experience, this theory can be salvaged, if at all, only by noting that parties make certain errors in the setting of the approximate standard of care, or, in the alternative, rely on escaping detection from the legal system.

77 Landes & Posner, The Positive Economic Theory of Tort Law, 15 Ga. L. Rev. 851, 880-83 (1981); Landes & Posner, Joint and Multiple Tortfeasors: An Economic Analysis, 9 J. Legal Stud. 517 (1980). Note that Landes and Posner do not explicitly consider the case of employer and employee. But if they are correct in the general view that the tort law mirrors the types of allocation schemes that parties would choose if transaction costs were low, then their theories should predict contracts that use negligence and contributory negligence to allocate risk, and not the compensation system.

Once this point is acknowledged, however, the entire theory breaks down, for it no longer can be claimed that both employer and employee will act on the assumption that the other side will not be negligent—or even erroneously not found to be such. We know it; they will know it too, both in the abstract and on the strength of their personal experience. Everywhere there can be evidence of negligence, whether in the condition of the equipment or in the design of the plant. At this point each party will no longer act as though he alone will be responsible for the harm; instead, each will take into account the possibility that the conduct of the other will be culpable and thereby reduce the costs of his negligence. Now both parties will, if rational, assume that they will have to bear only a part of the costs of their own negligence. The employer may choose, therefore, to take less than appropriate care on the expectation that the worker will be negligent; the worker may adopt the same attitude in the hope that he will escape detection and the employer will be found negligent. We have a non-cooperative strategic game in which optimal outcomes are hard to achieve. As each party never faces the full costs of his own neglect, the incentives under the common law rules can no longer be ideal.

We are thus back to the very situation upon which workers' compensation rests. There is no way to have ideal incentives on both parties at the same time so long as the system is so constrained that once the accident occurs, the set of possible outcomes is limited to transfer payments from employer to employee. It is, therefore, no objection to workers' compensation that it does not make both employers and employees face the full costs of the accidents they create, because this objection is valid against every conceivable alternative, including the negligence system. That this was understood intuitively before workers' compensation seems evident, as assumption of risk—in which cost-effectiveness is irrelevant—played a far greater place in workplace accidents than contributory negligence.\(^7\) For all their differences, the assumption of risk defense and the workers' compensation rules share the explicit rejection of the fine tuning through reasonableness necessary to make negligence rules work. In truth, reasonableness rules are too complicated to be attractive in any consensual situation. It is always more difficult to measure inputs than to determine outputs.

---

\(^7\) See Rabin, supr ante note 17, at 939, 947.
Landes and Posner do suggest that a decrease in information costs makes the evaluation of inputs cheaper today than it once was.\(^6^0\) Factually, this assumption seems quite wrong, as the procedures for adjudication are far more complex today than they were at the turn of the century, especially within the American context. Even if it were true, that assumption also makes it possible to monitor outputs more cheaply than before and cannot of itself generate a preference for the input-based negligence rule. Workers’ compensation rules are in most instances a closer approximation to the consensual ideal than the negligence rules to which they are opposed. The test of efficiency is what the market “would have” demanded were transaction costs low. By that standard negligence is an inefficient system, notwithstanding the flawed theoretical arguments offered in its support.

C. The Institutional Framework

The implicit logic of this critique of negligence and qualified defense of the compensation system is better revealed when the compensation rules are set against a broader background of contract and tort principles. Here it is necessary to discuss in some detail two separate doctrinal lines: first, the restrictive awards for damages in contract cases under the rule in *Hadley v. Baxendale*;\(^8^1\) and second, the operation of the general privity limitation as it applies to products liability actions. For these purposes, moreover, the nineteenth-century English and American law can be treated as a piece, as the differences between them are small and unimportant, notwithstanding the major divergence that has come to characterize the two systems in the twentieth century.

1. *Hadley v. Baxendale*. The rule in *Hadley v. Baxendale* marries the broad standard of strict liability for breach of contract to a narrow principle of damages. That point is evident on the facts of *Hadley* itself, as the plaintiff was denied the recovery for lost profits that resulted from the defendant carrier’s failure to deliver a needed shaft to the plaintiff’s mill. The linkage between stringent rules of liability and limited rules of damages has often been regarded as an oddity.\(^8^2\) In my view the two doctrines fit together

---

\(^8^1\) 156 Eng. Rep. 145 (Ex. 1854).
\(^8^2\) See Farnsworth, *Legal Remedies for Breach of Contract*, 70 COLUM. L. REV. 1145, 1216
hand in glove. In part the secret of their success lies in the discussion of the issue by Mayne in his classical treatment of damages:

[I]n matter of contract, the damages to which a party is liable for its breach ought to be in proportion to the benefit he is to receive from its performance. Now this benefit, the consideration for his promise, is always measured by the primary and intrinsic worth of the thing to be given for it, not by the ultimate profit which the party receiving it hopes to make when he has got it.83

There is, in a word, no presumption that either expectation or consequential damages are recoverable for breach of contract. In tort suits against strangers, they are appropriate; the defendant has internalized all the benefits, so too should he bear all the costs. But as benefits under a contract are shared, so too we can expect costs to be shared, as the parties will arrange their affairs in ways that maximize their joint profits and use price adjustments to match the residual risks assigned to each party.

With damages below losses, why then should the defendant be left a large and expensive "out" on liability? The strict liability rule clarifies the conditions of breach; the elimination of consequential damages reduces the premium needed by the defendant, and the built-in shortfall that falls upon the plaintiff gives him a stronger incentive to mitigate losses than any legal requirement to do so, with its attendant factual uncertainty. A combination of strict liability with limited damages acts like a settlement of a case before the breach, to be applied after its occurrence. The attractiveness of the general rule is confirmed by the numerous explicit contracts that reach the same general position—strict liability, liquidated damage clauses, no consequential damages, repair and replacement only—and that are today the subject of so much litigation.84

The point about Hadley v. Baxendale, and the regime of con-

---

tract damages it presupposes, is not a digression, but is central to
the explanation of the success of workers’ compensation systems.
There is in the literature a common refrain that insists that the
general principles of contractual freedom should not apply to cases
of personal injury: mistakes of information, misperception of the
risk, and the dominant position of the employer are all said to pre-
clude this happy development. Indeed it is typically said that the
compensation system is needed precisely to respond to these
problems. But why assume the worst? As the agreements in Griffi-
ths and Clements suggest, workers’ compensation is often the
market response to the accident question. In so doing, voluntary
systems of workers’ compensation adopt the very same strategy for
allocation of losses that Hadley v. Baxendale adopts in commercial
settings. The “arising out of and in the course of employment” re-
quirement is a strict liability provision, analogous to the strict defi-
nition of breach of contract. The limitation of damages to some
multiple of the weekly or annual wage is a prohibition against full
tort recovery parallel to the contractual prohibitions against conse-
quential damages or lost profits so routinely found in the negoti-
ated agreements. The subject matter of the risk may be impor-
tant—as human life is more important than lost profits. But it
does not follow under conditions of scarcity that the best possible
strategy for the minimization of loss must therefore change. No
perfect rule is possible, so the decisive question is what rule has
the fewest flaws. From this it follows that the great secret for suc-
cess of the workers’ compensation system lay not in its vaunted,
coercive original compulsion, but in the fact that it followed the
very pattern of risk distribution that both historical experience
and general theory of contract law indicated would best minimize
the risks in question. The repeated references to the workers’ com-

85 See, e.g., 1 A. Larson, supra note 5, § 2.20, where the compensation system is said to
rest upon a rejection of both “an individualistic moral code” and a system that provides
workers with general welfare payments in the form of “county relief” or a “direct handout.”
See also Henderson, Should Workmen’s Compensation Be Extended to Nonoccupational
Injuries?, 48 Tex. L. Rev. 117 (1969). The adoption of expansive coverage rules is accept-
able only if we view the compensation system as having the function of compensation after
the fact. When loss minimization is regarded as one goal of the system, that view must be
altered. Thus one reason why nonoccupational injuries are not covered is the distinct possi-
bility that individual workers with a conservative lifestyle will be forced to subsidize those
with a taste for risk. With the work-relatedness criterion in place, this implicit, but un-
wanted, transfer will be less likely to occur.
pensation bargain have more substance than rhetoric. As with the rule of common employment, the proposition should be qualified, for it may well be that the variation within the class of the employer-employee relationship is such that a single pattern of compensation does not work well in all instances. But if the structure of workers' compensation is best attuned to firms with dangerous employments, then the fit is apt to be best where it is most important. And if, as is usually the case, the statutes themselves are all brought into play with plants above a certain size, the fit is still improved, as those businesses that are too small to justify the fixed investment in the workers' compensation system are excluded from its operation. The contractual norm thus accounts for the regulatory success.

2. The Privity Limitation. The second of the major institutional features of the workers' compensation system is the privity limitation. This limitation provides that an injured plaintiff can sue a defendant responsible for the manufacture or sale of a defective product only if the two parties are directly bound to each other by contract. Let an intermediate party be introduced—e.g., a retailer between a manufacturer and consumer—and the plaintiff's case need not vanish in its entirety, but it automatically will be redirected against the party with whom the plaintiff has immediate contractual arrangements. The privity limitation has uniformly been jettisoned in the physical injury cases, even though it continues to retain some strength in cases of commercial loss. It is said that the privity limitation is, first, productive of massive injustice in that it insulates manufacturers from ultimate consumers and, second, a source of great administrative costs in that it requires an elaborate network of contract actions in which the ability to reach the ultimate defendant may be limited by insolvency, contractual defenses, statutes of limitations, or jurisdictional difficulties. The standard view on privity is so well entrenched that it

---

**86** See text accompanying notes 15-16 supra.


**88** For the limitation's original statement, see Winterbottom v. Wright, 152 Eng. Rep. 402 (Ex. 1842).

**89** See Restatement (Second) of Torts § 402A (1965).

**90** See, e.g., Evra Corp. v. Swiss Bank Corp., 673 F.2d 951 (7th Cir. 1982).

**91** See Prosser, The Assault Upon the Citadel (Strict Liability to the Consumer), 69
might seem idle to say a kind word in its behalf, but in order to understand the successful operation of the workers’ compensation system in its formative years, it is necessary to do just that.

The general attack on the privity limitation rests on an important error of undergeneralization. The typical cases in which the issue is raised involve ordinary consumer goods sold at retail. This will include the famous belladonna of Thomas v. Winchester,\textsuperscript{62} or the defective wooden wheel of MacPherson v. Buick Motor Co.\textsuperscript{63} Within this context the privity doctrine is quite unappealing because the immediate seller is typically a passive conduit of a product that is designed and prepared by others. Why, therefore, should the soundness of the product be litigated in a suit brought against the defendant who is least responsible for its merits? It is just another variation on the familiar theme that we do not shoot the messenger because he brings the bad news.

For these purposes we can assume that this limited critique of the privity doctrine is wholly sound.\textsuperscript{64} The important question is how well does it carry over to the ordinary workplace accident. Here the key point of difference is that the employer is a very different intermediary from a wholesaler or even a retailer. For one thing, the employer himself is always within the jurisdiction; for another, he has had and retains direct and immediate control over all that is used upon his premises. It is the rare machine tool that is simply uncrated and placed into operation without customization or maintenance. It is the rare employer who is a mere conduit of goods produced by others. Within this setting there is an enormous advantage in concentrating the litigation over industrial accidents against the employer. There is but a single defendant in the ordinary action. The elimination of peripheral parties reduces administrative costs dramatically. Targeting the action against the employer reorients the structure of the basic action on a host of important substantive issues. All the modern difficulties in establishing the identity of a supplier of a defective product or component part are gone.\textsuperscript{65} The questions of causal intervention and

\textsuperscript{62} 6 N.Y. 397 (1852).
\textsuperscript{63} 217 N.Y. 382, 111 N.E. 1050 (1916).
\textsuperscript{64} For a more extended statement of my own views, which does not address the point made here, see R. Epstein, supra note 15, at 9-24.
\textsuperscript{65} See, e.g., Sindell v. Abbott Laboratories, 26 Cal. 3d 588, 607 P.2d 924, 163 Cal. Rptr.
product modification, ever recurrent in modern product liability actions, are virtually eliminated by the absence of intermediate parties. Whatever the appropriate liability regime for industrial accidents, the system operates better when the proper defendant is the immediate employer and not some remote party to the case. Thus the privity rule makes sense no matter whether the relationship between employer and employee is governed by common law negligence or compensation rules. If we could ask what network of contracts between these parties would emerge if transaction costs were zero, a good first approximation would be the corner solution—no liability of the remote party to the plaintiff. Actions over under express agreement might take place, but I should regard them as unlikely occurrences. The spirit of Hadley v. Baxendale would penetrate to this important corner of the law: limited damage rights under contract, coupled with inspection and servicing of equipment, is the more likely approach.

If this is correct, then the success of the old system depended in large part upon the existence of the privity limitation on the third-party actions against product suppliers of all stripes and conditions. With those direct actions barred, the possibility of third-party actions against the employer could not arise, thereby preserving the structural integrity and simplicity of the workers' compensation system. The allure of expanded torts damages is always present, but to maintain the structure of the basic compensation system that allure had to be resisted not only by the parties, but also by the judges.

IV. Degeneration, and Revival?

This Article passes over large portions of the history of workers' compensation to address this single question: what accounts for the major challenges to the workers' compensation system today? In order to answer this question, it is necessary to stress not the strength but the one major weakness of the original legislative bargain: it was rigid and not capable of self-correction. Here, I shall not detail the various shifts that have taken place within the compensation system but shall only point to them in the most general

132 (1980).
** The modern source of these actions is found in the important case of Dole v. Dow Chem. Co., 30 N.Y.2d 143, 282 N.E.2d 288, 331 N.Y.S.2d 382 (1972).
The central point is that both the definition of what counts as a compensable event and the levels of compensation awarded have been expanded through judicial interpretation and legislative enactment in recent years. A priori, it is difficult to say whether these changes are a good or a bad thing without knowing whether they move with or against the dominant contractual norm. My own suspicion, judging from the present distress of the system, is that the levels of benefits and the extent of coverage have simply become too lavish to satisfy the elaborate set of constraints upon any compensation system, voluntary or compulsory. The large protests from firms required to pay them—coupled with oft realized threats either to leave the jurisdiction or to locate expansion elsewhere—are consistent with this view.

There is, however, an incipient development that, if it obtains general acceptance, poses a greater threat to the viability of workers' compensation. It concerns the various judicial efforts to dismantle the exclusive remedy protection that the system affords the employer. I wish to concentrate only on one aspect of the problem—the direct tort actions brought by the employee.

There are two recent developments that require some special comment. The first of these is the radical expansion of the "exclusive remedy" protection. This was illustrated by the Illinois legislature's sharp increase in the benefit levels under the compensation system in what appeared to be an agreement between then Governor Walker and leading labor leaders. The response to the increases was sharp, as many business and insurance leaders complained that Illinois had become a state bad for business. See, for various stages of the controversy, Houston, New state law on injuries at work costing millions, Chi. Tribune, Feb. 22, 1976, § 2, at 13, col. 1; Wiedrich, State laws make city's business climate foul, Chi. Tribune, Jan. 13, 1980, § 1, at 5, col. 4.

The proposal developed in this paper did not go the full length of blocking all third-party actions, although there is, as I have said, much to favor that position. Instead, it sought to coordinate better the interaction of the compensation and tort systems by (1) reducing the amount of tort recovery against a third party by the compensation benefits paid or payable, (2) eliminating the rights, be it by lien, subrogation, or direct action, of the employer to recoup his compensation benefits from the tort recovery of the employee, and (3) blocking all actions for indemnity or contribution brought by the third party against the employer, except by express agreement. The advantages of this proposal are chiefly two. First, it reduces the administrative costs of operating the system; and second, it prevents the limited financial incentive upon the employer from being dissipated, as it is when he recovers, regardless of his own negligence, his full compensation payments from the third-party defendant. Note that in Epstein, supra, I addressed this scheme only in the context of a third-party products liability action. Its application seems, however, more general in scope.

---

9 In Illinois, for example, the legislature sharply raised the benefit levels under the compensation system in what appeared to be an agreement between then Governor Walker and leading labor leaders. The response to the increases was sharp, as many business and insurance leaders complained that Illinois had become a state bad for business. See, for various stages of the controversy, Houston, New state law on injuries at work costing millions, Chi. Tribune, Feb. 22, 1976, § 2, at 13, col. 1; Wiedrich, State laws make city's business climate foul, Chi. Tribune, Jan. 13, 1980, § 1, at 5, col. 4.

** See Epstein, Coordination of Workers' Compensation Benefits with Tort Damage Awards, 13 Forum 464 (1978). The proposal developed in this paper did not go the full length of blocking all third-party actions, although there is, as I have said, much to favor that position. Instead, it sought to coordinate better the interaction of the compensation and tort systems by (1) reducing the amount of tort recovery against a third party by the compensation benefits paid or payable, (2) eliminating the rights, be it by lien, subrogation, or direct action, of the employer to recoup his compensation benefits from the tort recovery of the employee, and (3) blocking all actions for indemnity or contribution brought by the third party against the employer, except by express agreement. The advantages of this proposal are chiefly two. First, it reduces the administrative costs of operating the system; and second, it prevents the limited financial incentive upon the employer from being dissipated, as it is when he recovers, regardless of his own negligence, his full compensation payments from the third-party defendant. Note that in Epstein, supra, I addressed this scheme only in the context of a third-party products liability action. Its application seems, however, more general in scope.
dual capacity doctrine. The second deals with the parallel expansion of the employer's liability for intentional harms.

A. Dual Capacity

That some case for a "dual capacity" doctrine exists is certain. Suppose that on Sunday, A runs down B on a public street while driving his own car. Should B's action in tort be barred because on another day and in another place B turns out to be the employee of A? The explicit premise of the legislative bargain is that the compensation system extends only to those accidents that arise out of the employment relation. It is, in a word, a relationship not merely between persons, but between persons as holders of well-defined roles. To admit the extreme case, however, carries with it the cost of line drawing in more marginal situations. In *Duprey v. Shane*, the plaintiff was a nurse in the employ of the defendant chiropractor. The nurse was injured in the course of her employment, and the defendant physician, who carried compensation insurance, treated her on the same footing as his other patients. The court held that the action for medical malpractice could proceed, notwithstanding the exclusive-remedy provision of the compensation law, and argued that in treating his nurse the defendant had "assumed a role distinct in both time and nature from that of an employer." In my view the case is tenuous, if not wrong, for the plaintiff did not come to the defendant qua patient with an injury wholly unrelated to her employment. This should certainly be the case where medical care is provided as part of the ordinary course of the employer's business, without any additional charge to the employee. In the typical case of on-the-job medical treatment and even in *Duprey* itself, it seems better to regard the treatment as a continuation of the employment relationship that is accordingly covered by workers' compensation.

To be sure, the integrity of the compensation system could survive if the erosion of the exclusive-remedy provision were limited to the occasional case where an individual employer was treating his own employee. Matters, however, take on a more ominous

---


*101* See 2A A. Larson, supra note 5, § 72.61(c) (urging that the doctrine be "confined to
appearance if the dual capacity doctrine can routinely overwhelm the exclusive-remedy defense in all employment cases. The linguistic trick behind this massive institutional transformation is easy to see. Dual capacity properly applies when $A$ and $B$ are brought together only in some relation wholly separate from that of employer and employee. This is worlds apart from the case where they are brought together in some relationship in addition to that of employer and employee.

Just this fatal step was taken in the ill-considered case of Bell v. Industrial Vangas, Inc.\textsuperscript{102} There the plaintiff was a route salesman for the defendant and was responsible for the delivery of natural gas to the defendant's various customers. The plaintiff was injured by a fire that occurred during the course of one of those deliveries. There could be no serious doubt that plaintiff was entitled to workers' compensation. The only question was whether the benefits so recovered represented the ceiling, or only the floor, of his rights against his employer. To convert the system into a floor only, the plaintiff alleged that the defendant engaged in the manufacture, packaging, and selling of the natural gas in question and therefore was covered by the California law of products liability, which concededly would have applied if the plaintiff had not been an employee of the company. The case therefore raised anew the fundamental challenge of the integration of two separate systems of liability, each with its own history and doctrinal base.

The California Supreme Court held that the products liability rules, whose policies of accident protection and loss distribution are so antithetical to the workers' compensation system, applied. It therefore held that the occurrence of the accident within the course of employment was only a "mere fortuitous circumstance"\textsuperscript{103} that did not in and of itself bar a products liability action. Yet the circumstance was not fortuitous. That the plaintiff

\textsuperscript{102} 30 Cal. 3d 268, 637 P.2d 266, 179 Cal. Rptr. 30 (1981).

\textsuperscript{103} Id. at 279, 637 P.2d at 273, 179 Cal. Rptr. at 37.
was an employee keyed him into the compensation system in a way the rest of the world was not. As the decision itself did nothing to remove the employee's compensation benefits, it did not restore any implicit parity between workmen and strangers. Nor is there any reason, if this were possible, for it to be done. Whatever the weakness of the common-employment doctrine, it recognized, as does the workers' compensation law, that employees and strangers fall into separate categories. Where an explicit bargain limits B's rights of action against A, it should be respected even if C's rights of action against A remain wholly unimpaired. A contract individualizes treatment; it takes parties out of the mass of strangers. No agreement could have any institutional function if it was valid only if it confirmed the status quo. The legislative bargain should be treated in the same fashion: it binds the parties to it. So long as workers and strangers are treated differently under the compensation system, there is no reason to insist that they should be treated in the same fashion under the tort system.

Yet in essence this is what the decision in Bell did. It takes little imagination to see that a bit of ingenuity can extend its rationale to cover not only the products liability actions, but also all other forms of tort suit. The dissent in Bell was without question correct in observing that “[i]f an employer is to be held civilly liable to injured workers in the employer's capacity as a 'manufacturer,' what compelling reason can exist for denying similar liability for injuries attributable to the employer's other relationships including his status as 'landowner,' 'motor vehicle operator,' or 'cafeteria proprietor'?"104 The decision in Bell can be criticized forcefully on the ground that it runs roughshod over the statutory language and therefore judicially repeals the basic legislative design. Yet given our analysis of the economic structure of the compensation laws, the decision would be equally ill-conceived if introduced by an explicit legislative enactment. The principles of economic scarcity are still at work. There is not enough value received from the individual worker to support two compensation systems. The point bears especial force when it is noted that the negligence law of today is vastly more favorable to the plaintiff than it was on the eve of workers' compensation. Fellow servant is gone; contributory negligence is comparative negligence; assumption of risk is now part of

104 Id. at 287, 637 P.2d at 278, 179 Cal. Rptr. at 42 (Richardson, J., dissenting).
comparative negligence. It is difficult to imagine any serious action in which the plaintiff will be unable to reach the jury against the employer if the exclusive remedy is not allowed to stand, as errors in plant design, supervision, warnings, or whatever will always be lurking in the background, even if not fully provable. One might argue that these developments in the law of negligence are welcome improvements over the earlier law. Yet here there is no evidence that the modern negligence doctrines bear any relationship to the legal rules that would be generated by voluntary transactions and every evidence that the courts that impose them would not tolerate their removal by contractual means. The private power to correct for judicial errors has been removed. At any rate, a system that contemplates both the tort action against the employer and the compensation is beyond rational defense.

It is in one sense too early to determine the institutional effects of Bell and kindred cases. In California the decision itself has already proved to be a prelude to further legislation, for a recent statute overturns Bell, increases the overall level of compensation, and creates, for reasons that appear wholly unprincipled, an exception to the exclusive-remedy provision for those machine tools that are improperly installed or maintained by the employer. The three parts of the bill have no organic unity. It is simply a reflection of how judicial trumps can be transformed into legislative benefits. Outside of California, there have been many decisions that have decided the question in opposite ways, and even in those jurisdictions where dual capacity has been expanded, subsequent judicial decisions could again limit its scope. If not, then at the very least some corrective legislation (that does not follow the California pattern) is needed to restore the original status quo ante. It will not do to have one side of the legislative bargain undone with the other still in force. Nor will it do to remove both sides of the bargain. What is needed is a return to the original conception of the system with its robust exclusive-remedy provision.

B. Intentional Torts

The second direct challenge to the exclusive remedy provision is

---

106 See Bell and cases cited therein, 30 Cal. 3d at 287, 637 P.2d at 278, 179 Cal. Rptr. at 42.
found in the treatment of willful or intentional wrongs by the employer. Unlike ordinary negligence, intentional harms introduce an element of moral hazard that is very difficult to control by a set of rules designed for accidents. This point is, I think, confirmed by the general treatment that deliberate wrongs have received in other contexts. For example, a defendant who commits an intentional harm cannot set up contributory negligence or plaintiff's misconduct as a defense, and the compensation law similarly recognizes the problem by barring recovery for a worker in willful default.

How then should the employer's intentional wrong be treated? One possibility is to permit a direct action in the tort system, as if the compensation laws had never been enacted. But this approach is, I think, inferior to the alternative that, within the framework of the compensation system, makes the employer pay some penalty (say 150% of normal benefits) to employees that he intentionally injures. The advantages of this approach are two. First, it keeps the entire case within the workers' compensation system, so that there is no need to worry about the coordination of tort and compensation actions. Second, it protects the system against the indirect erosion that occurs when the intent requirement is attenuated to include mental states that are insufficient to support either criminal punishment or (in a world more sensible than our own) punitive damages.

These possibilities for direct tort actions against the employer have borne fruit within the past several years, as recent cases have shown a marked willingness to allow cumulative trauma actions to be pursued against employers outside the workers' compensation system on the ground that the injuries are intentionally inflicted. The allegations in question are typically of two particular sorts. The first and more general claim is that the deliberate exposure to dangerous substances in the workplace, especially if in willful violation of statute, supports a tort action for both actual and punitive damages. The second, and narrower, position is that the improper provision of medical inspection and treatment after the original exposure, if "falsely and fraudulently done," supports the tort actions, again for actual and punitive damages, but only if the deliberate conduct aggravated the original injury. The first of these

---

107 For my views, see Epstein, Intentional Harms, 4 J. LEGAL STUD. 391 (1975).
positions has just been adopted by the Ohio Supreme Court in Blankenship v. Cincinnati Milacron Chemicals Inc. The second of these propositions was accepted by the California Supreme Court in Johns-Manville Products Corp. v. Contra Costa Superior Court (hereinafter Rudkin, the claimant), which, however, expressly rejected the broader proposition subsequently embraced in Blankenship. Both of these breaches of the exclusive remedy provision should be rejected. This is clear enough on statutory grounds, as the language was drafted to guard against this very result. The statute to one side, the exclusive-remedy provision is critical to the basic legal structure for industrial accidents, especially in an environment in which private agreements are unenforceable.

In embracing the tort remedy, the Ohio court in Blankenship started with the premise that the basic compensation bargain prevented the worker from maintaining tort actions for negligent harms but not for those intentionally inflicted. In making this particular opposition, however, it never offered a coherent account of what constituted an intentional wrong, although the concept is far from self-evident. With an ordinary traumatic injury a deliberate assault by the employer upon a worker is outside the coverage of the workers' compensation statute, given the specific intention to harm and the utter want of any justification or excuse for the employer's conduct. By the same token, however, an ordinary personal injury case remains within the compensation system where the employer knows, even to a certainty, that equipment will necessarily cause harm to some workers, perhaps because it will be defective, or ill-maintained, or housed in violation of statute. What is at stake in all these cases is the looser form of intent—the mere knowledge that future injury will take place as the by-product of productive activity. One justification for the introduction of the workers' compensation statutes was that they instituted a certain predictable response for the "inevitable toll" of all productive activity. To take these cases out of the compensation system is to dismantle it before it is scarcely in place; to my knowledge, at least, no court has required that result.

Should the rules of the game change when we move from ordi-

---

106 69 Ohio St. 2d 608, 433 N.E.2d 572 (1982).
nary accident cases to cumulative trauma cases? There is no question but that the two sorts of cases are quite different. The suddenness is gone with cumulative trauma, and with it the ability to locate the source of the accident at a single time and place. In the early days of workers' compensation, the problem of dealing with cumulative trauma caused by diffuse and multiple sources was regarded as so great that the cases were held as falling outside the coverage language—"personal injury by accident"—unless it could be shown on the model of ordinary accidents that a discrete and sudden exposure was responsible for the identified harm. Today, most statutes have special provisions to cover occupational diseases. But these continue to present special difficulties, especially with workers who have changed jobs many times during their career. Here, it is difficult to identify the proper employer for any given injury or disease, even assuming the condition itself to be work related. And it is therefore wholly proper to adopt very strict administrative rules to handle the coverage question, such as the rule that locates each cumulative trauma case to the last insurer of the last employer responsible for the condition in question. But the differences created by cumulative trauma cases do not require any other modification of the compensation system. Whatever definition of the intention was proper with ordinary personal injuries should carry over to this new context as well. It should therefore be wholly immaterial that the employer knew that injuries could result from exposure to substances found in the workplace or even that exposure levels were higher than permitted by statute. Knowledge of the first form is always present, and breaches of regulatory duties within the compensation framework should be the source of statutory fines and penalties only. These breaches should not have the enormous consequence of taking a case out of the compensation system itself, by creating a private right of action in tort. It follows, therefore, that only intentional wrongs, in the narrow sense of deliberate infliction of harm for its own sake, can allow tort actions against employers in the cumulative trauma area, a rare and institutionally unimportant situation.

The California court in Rudkin recognized the force of the arguments against the position in Blankenship, but nonetheless insisted that the provision of the medical inspection program offered

---

the plaintiffs a possible avenue of escape from the compensation system. Again, however, the parallels to ordinary trauma cases should be decisive. The question of medical treatment of injured workers arises frequently in connection with ordinary accident cases. Yet there is no reason to allow malpractice actions to be routinely brought with respect to such treatment. Instead, the proper line is to insist that the tort action be allowed only when there is a deliberate infliction of harm for its own sake against the worker, a case that will typically require proof of actual malice on the part of the treating physician. The same logic again applies with cumulative trauma cases. Wherever the inspections are conducted on a voluntary basis, it seems very odd to tell an employer who has reduced the expected levels of injury that he is now to be held to full tort damages for the injuries that do in fact occur. A decision of that sort can only penalize desirable actions and therefore make it more likely that those actions will not be undertaken, to the detriment of the very workers whom the statute is designed to protect.

The Court in *Rudkin* said, and with some conviction, that the compensation statutes should not serve as a shield for palpable misconduct. But the response is that, unless the statute provides that kind of a shield in all but the most extreme cases, the compensation system will lose its effectiveness in industrial accident cases. The very fact that the plaintiffs are allowed the limited escape in *Rudkin* now means that both parties will have to incur the very heavy expenses of litigation in order to determine whether the limited cause of action stated by the plaintiff can be proved. All company officials and physicians who participated in the inspection plan become fair targets for the depositions and interrogatories that so disrupt the ordinary management of the firm. The causal issues will be difficult because the plaintiffs cannot recover if the original exposures were sufficient to cause harm. And if the defendants prevail on the merits, it means that they will have

---

111 The similar problem arises in all cases of immunity. See, for example, Harlow v. Fitzgerald, 102 S. Ct. 2727 (1982), where the Supreme Court adopted a qualified privilege doctrine for the confidential advisers of the President, while in the same breath enjoined lower courts to issue summary judgments for the defendant in order to stave off the possibility of extensive litigation.

112 On this point *Rudkin* goes no small way toward easing the plaintiff's burden by noting that where it is impossible to apportion damages between those exposures that could have been avoided with proper medical care and those that could not, the presumption should be
been required to bear the costs of these complicated lawsuits without any protection at all. The goal of the compensation statutes was that uncertain remedies should be replaced by certain ones, so as to prevent litigation from becoming a grotesque imitation of global war. Those gains to compensation can be obtained only if the temptation is resisted to look behind the veil in individual cases. The exception for intentional harms must be limited to its narrowest possible scope, especially if it is sufficient to take cases out of the compensation system and even if the plaintiff receives an increased recovery within it. I have little doubt that under the compensation systems developed by voluntary means these escape provisions would not be expanded to permit suits like Rudkin to go forward, while the outcome in Blankenship is wholly inconceivable as a contractual matter. The courts that are now engaged in the wholesale attack upon the exclusive-remedy provision make all too little of its institutional importance.

How then can the compensation system insulate itself from the shocks of these similar decisions? The instinctive reaction is for corrective legislation designed to shore up the exclusive-remedy provision. But the current onslaught might well spur us to reconsider the foundations of the entire system. Why does workers' compensation have to rest upon a compulsory legislative action? Why is it not possible for it to be the outgrowth of a voluntary arrangement? I would not be troubled with a statute that said that in the absence of express agreement compensation benefits stipulated by statute were awarded; that in itself would represent an enormous change from the older common law. But the statutory schedules could be the beginning and not the end of the process. If the legislature set the coverage formula incorrectly, it could be changed by private agreement. And the same would be true of benefit levels. In many cases if the system as a whole proved unattractive, contracting out could be complete.\(^{113}\) I do not think that we should be troubled by this result if it is arrived at by voluntary processes. Indeed, in an age of collective bargaining, there seems no reason why, in the typical manufacturing context, negotiations

---

over compensation cannot take place side by side with those over wages, working conditions, and fringe benefits. Similarly, employers have incentives to supply the right mix of benefits even for a nonunionized workforce, which typically receives all forms of health insurance, disability pay, and life insurance. There are, of course, questions as to whether workers would be duped or deceived, but I think this highly unlikely in any broadscale and durable institutional arrangement. In truth, the risks running in the opposite direction may be far greater today. If legislatures and courts forget the historical origins and economic structure of the compensation statutes, they may bring about an institutional disarray that works against the interests of both the employer and the employee.