RESTORING SANCTITY OF CONTRACT

IN EMPLOYMENT RELATIONSHIPS

NEW ZEALAND BUSINESS ROUNDTABLE
NOVEMBER 1999

Richard A Epstein

This lecture, Restoring Sanctity of Contract in Employment Relationships, was given in Wellington on 24 March 1999 at the offices of Chapman Tripp Sheffield Young.

First published in 1999 by
New Zealand Business Roundtable,
PO Box 10–147, The Terrace,
Wellington, New Zealand
http://www.nzbr.org.nz

ISBN 1–877148–57–1

Cover illustration by *Jo Tronc, Watermark Ltd, Wellington*
Design and production by *Daphne Brasell Associates Ltd, Wellington*
Typeset by *Chris Judd, Auckland*
Printed by *Astra Print Ltd, Wellington*

Contents

Richard A Epstein

RICHARD A EPSTEIN is the James Parker Hall Distinguished Service Professor of Law at the University of Chicago, where he has taught since 1972. Previously, he taught law at the University of Southern California from 1968 to 1972.

He has been a member of the American Academy of Arts and Sciences since 1985 and a Senior Fellow of the Center for Clinical Medical Ethics at the University of Chicago Medical School. He served as editor of the *Journal of Legal Studies* from 1981 to 1991, and since 1991 has been an editor of the *Journal of Law and Economics*.

His books include *Torts* (Aspen Law and Business, 1999), *Principles for a Free Society* (Perseus Books, 1998), *Mortal Peril: Our Inalienable Right to Health Care?* (Addison Wesley, 1997), *Simple Rules for a Complex World* (Harvard, 1995), *Bargaining with the State* (Princeton, 1993), *Forbidden Grounds: The Case Against Employment Discrimination Laws* (Harvard, 1992), *Cases and Materials on Torts* (Little, Brown, 5th ed, 1990), *Takings: Private Property and the Power of Eminent Domain* (Harvard, 1985), and *Modern Products Liability Law* (Greenwood Press, 1980).

Professor Epstein has written numerous articles on a wide range of legal and interdisciplinary subjects and taught courses in contracts, criminal law, health law and policy, legal history, property, real estate development and finance, jurisprudence and taxation, torts, and workers' compensation.

Restoring Sanctity of Contract in Employment Relationships

I can think of two possible approaches to the general question of how to restore the sanctity of property and contract – a systematic exposition from first principles or a more autobiographical account. On this occasion, I will begin with some personal reflections. Thereafter, I shall adopt a more theoretical stance to analyse the key issues relating to the sanctity of property and contract from the perspective of one particular set of disputes – employment contracts. My primary focus will be on contracts involving relationships with unions. Only toward the end will I move on to more modern employment issues such as unjust dismissal laws. These topics raise many of the same problems encountered in dealing with unions, but present some new issues as well. Because I have already discussed the contract at will several times in New Zealand, on this occasion I do not wish to address that topic again directly, although, of necessity, it will figure in the larger discussion.

The disguised virtues of an English legal education

I am first and foremost an English-trained lawyer. I received my basic legal education at Oxford from 1964 to 1966. It was an unusual education, both informative and formative. It was one that insulated me from certain

American outlooks on legal education that I would encounter shortly thereafter in 1966, when I returned to the Yale Law School. An English legal education had one distinctive characteristic that worked greatly in my favour. At Oxford we took grand subjects such as property, contracts and torts in all their forms. We did not bother with any subdivisions within a field. Everything was treated as part of a comprehensive general subject, which in turn made it easy to think of it as all part of some comprehensive theory. Thus in contract law the same basic principles applied to labour contracts as to contracts for the sale of goods, or indeed of any other type. We learnt the field as a whole rather than a series of specific sections, each resting on its own intellectual bottom.

Moreover, my English instruction – at least in 1964 – took place in blissful ignorance of the social fact that the administrative state was already upon us. We studied contract as a private law subject, concentrating on the opinions of the classical liberal judges of the late nineteenth century. Contract law for me was Baron Bramwell, Jessell MR, Lord Bowen and Judge Blackburn. The powers of the administrative state, which altered the way in which business was conducted in so many areas, were not included in the study of law because it lacked the appropriate theoretical aspirations. In consequence, by learning an idle theoretical subject with comparatively little practical relevance – the standard Oxford education – I learned far more than if I had been subjected to the sordid details of administrative law, where I might have discovered that the sanctity of property and contract were held in very low esteem. Thus I became an expert at barring the entail in the fashion of the great medieval land lawyers, procedure as obsolete in the twentieth century as it might have been useful in the thirteenth century. Yet never during my entire time in England did I learn about the operation of the Town and Country Planning Act or its consequences on the patterns and efficiency of land use in England. The one course we had in administrative law was far more concerned with the finality of agency determinations than with the substantive content of any administrative order.

This strange inversion of the theoretical and practical might have proved a major disadvantage to the English who aspired to practise law in their own country. But for me, an American, it was an unintended inoculation and inspiration that has guided me for the rest of my professional career. Evidently, judging from the outcome, there are dangers in allowing students to master law through independent study. They may be capable of learning for themselves. They may even become contrarians – not only when abroad, but also on returning home. I thus derive much benefit from my English education in ways that might have shocked or amused my teachers. I hold it in great affection, not solely because of what my teachers wished specifically to teach me, but rather from what they let me, and encouraged me to, teach myself.

Learning the law of contract, property and tort in a very broad sweep was not the only hidden advantage. At Oxford we also learned the subjects in a completely non-contextual manner. Whether the topic was offer and acceptance, or consideration, or damages, the only two people involved in a given example were a person named A and a person named B. These characters were truly ubiquitous, and if they owned property it was called Blackacre or something similarly non-contextual. There was no individuation. To a modern American contextualist, this approach shows a scandalous lack of sophistication about the realities of social and economic relations. With only A or B associated with a contractual relationship, and in the absence of special knowledge of the significance of the alphabet, one is unable to discover which party is 'dominating' the other. If only through desperation or ignorance, one must assume some parity of interest between the two parties – a parity which should be interrupted, if at all, only by the presence of C, who, being a latecomer, is typically at a legal disadvantage. This approach is immensely liberating, since it forces us to focus on the essentials of a contractual relationship. We learned a general theory rather than coming to imagine that contract law is a set of excuses put forward to explain why one party does not have the capacity to be bound by their undertakings.

There was one weak point in my English education. My teachers were excellent at explaining legal rules at the appropriate level of generalisation, and showing no fear or favour to either party. But they omitted a very important factor – why anyone would choose to enter into a contract at all. Contracts were seen as expressions and manifestations of the will. Promising was an act by which individuals bound themselves. There was no analysis of why anyone would undertake such an exercise, or why the courts would regard it as important to spend public resources on enforcing contracts. At that point in my education I needed to return to America. Eventually I came to understand why the structure I learned in England made much more sense than had been apparent to me when I took my final examinations at Oxford. The key is what I have termed the eleventh commandment. Every morning when I wake up, I face east, genuflect cautiously, and recite this commandment, to remind myself of my intellectual and spiritual orientation.

The eleventh commandment is that all voluntary exchanges benefit both parties to the exchange. Unless one grasps this insight, the entire subject of contractual relationships seems a mystery at best and a sheer absurdity at worst. Once understood, the eleventh commandment provides a powerful line of reasoning. If, as a general proposition, what is good for Party A can also be good for Party B, one need not inquire further into the role and motivations of the respective parties to understand why an exchange between A and B takes place. The mere existence of voluntary agreement between those parties provides a warrant that such transactions will generally be beneficial in the long run. When this is understood, suddenly the generalisations learned in England acquire a philosophical universality with great power to explain the whys and the wherefores of legal relationships.

The common law of an employment relationship

To illustrate how this general argument remains valid over more narrowly defined circumstances, consider labour contracts. If A and B are now identified as employer and employee, should that change one's view of the transaction? I will briefly discuss a few cases and statutes, some English and some American, to indicate the way the balance of power between the parties to employment relationships ebbs and flows. The first example is not in fact a contract case, but a tort case predicated on contract. This allows tort and property law to enter the analysis, enabling us to justify contract, tort and property law as an integrated structure.

The case is *Lumley v Guy*, decided in 1853, and it concerns inducement of breach of contract. This case stood for the principle that if an employer enters into a term contract with an independent contractor, in this case a famous opera singer, and a third party attempts to interfere with that arrangement by stealing the worker, the employer has an action in tort allowing them to enjoin the efforts of the stranger to lure their employee away, or to obtain damages if that interference has proved successful. Such a transaction is a simple A, B, C relationship, and possesses a commendable logic. In effect, before any contract is agreed between A and B, each party is entitled to metaphorically fish and trawl through open waters in search of a desirable contractual partner. In a world where no binding contracts have yet been formed, a full array of choice is available to both sides; that is what economists mean today by a competitive market in labour. But once one specific arrangement is agreed, and a contract is formed, those two parties have left the competitive sphere. Thus it is competition which leads people to make contracts – contracts that create exclusive rights between the parties. But the competitive market cannot work if those contracts are literally not worth the paper that they are written on. Some protection to the contract must be provided.

One possible remedy to the individual is protection against breach through a damage action, or perhaps an injunction: indeed the injunction was issued against Miss Wagner in the companion case of *Lumley v Wagner* (1852). The importance of remedies of this type should not be underrated. But for reasons discussed later, there are questions about the adequacy of these remedies. In consequence, in order to protect the contractual relationship, often attention focuses on Party C with a view to preventing that person from stealing the employee in the first place. One can obtain an anticipatory remedy, and need not concern oneself with the problematic business of damages after a relationship has been disrupted.

There is also the problem of how one accommodates C. If C is trawling through the waters, and spies an apparently unattached B, C may imagine that B is a potential contractual partner in an ordinary competitive setting. Without really knowing it, the common law, in effect, imposes on the parties the standard requirements of good faith contracting developed in connection with multiple transactions over the same piece of property. If C happens to encounter B, and has no notice of B's prior relationship with A, C can enter into a contract with B in the confidence that, if a Party A turns up, A's sole remedy will be against B. One cannot in effect impose a property right over an individual (ie Party B) that is secret to the rest of the world, any more than one can impose a secret lien good against the world on a piece of property. But if the third party does have *notice* of the relationship, that notice serves as a massive 'keep off' sign. To the extent that C now contracts with B in a way that adversely affects A, C is fishing in forbidden waters and can be sued. Given the disruption of the contractual relationship, C could be liable for damages.

Yet no matter how contract damages are calculated and enforced, they typically will be inadequate. There are inefficiencies associated with the various procedural matters involved in litigation. And with 'consequential damages' – those dislocations that flow from the breach – there is always the annoying possibility of being unable to prove with sufficient clarity what would have happened in the absence of the breach of contract. To switch metaphors from sea to land, the smart entrepreneur prefers to keep a given train on the tracks rather than attempt to calculate how far it would have travelled had the derailing never occurred. In other words, the preferred remedy starts from a presumption of wanting no breach to take place at all and provides an injunction against the third party, who cannot tamper with the relationship. Any contracting party, such as the opera singer in *Lumley*, might still choose to breach. But the probability of that occurring will be sharply reduced when there is no alternative avenue for employment.

The rules outlined here implicitly reject a modern doctrine of contract law that constitutes a mischievous misapplication of the general law and economics approach. That is the doctrine of the so-called 'efficient breach' of contract, an oxymoron that carries its own moral lesson.

The starting point for the doctrine of efficient breach is the assumption that all the original contracts are mutually beneficial to the parties concerned. So far so good. Suppose, however, a change in circumstances then leads one party to a contract (Party A) to have second thoughts about the benefits derived from the pre-existing contract. It is argued that Party A should be entitled to breach the contract if it pays Party B a sum sufficient to leave it indifferent as to whether or not the contract is breached. The economic logic is plain enough. Party B is no worse off in consequence of the breach, since the damages received make it indifferent as to whether performance under the contract goes forward. Party A is better off: it prefers to breach the contract even after paying compensation. And if one party is made better off while the other is no worse off, this is a so-called Pareto improvement that lands us happily in some economic Garden of Eden.

The problem with this reasoning lies with the practicalities of enforcing contractual remedies. The innocent party must bear the burden of suit, yet has no guarantee that it will obtain full and complete damages given the elusive nature of consequential losses. It is one thing to have a contractual remedy when individuals are simply unable to perform their promises, so that the affected party must be compensated. But to the extent that people engage in deliberate breach of contract, the insecurity felt by all individuals about their contractual arrangements will lead to a disturbing deterioration in general commercial relations. This is confirmed by trade association norms. Associations are always willing to contemplate expectation measures of damages when somebody is genuinely disabled from performing a contract. But they typically expel business people from their organisations if they engage in purely opportunistic breaches, in which they abandon one party and sell the same goods to another party. The behaviour that the doctrine of efficient breach extols, the world of commerce condemns. We can simply have no confidence that the damages remedy has the real world properties required by the doctrine of efficient breach.

Thus the common law comes to an extremely clever accommodation of the multiple interests involved in contractual relations by combining the notice provisions in property law with injunctive relief, reserving private damage actions as a back-up. The common law judges manage to obtain these right answers simply by doing their ABCs, with no explicit acknowledgment of the complex economic factors that impinge on the analysis.

Unions and the yellow dog contract

The importance of this ABC approach becomes even more apparent when one introduces many more people into the mix – parties D, E, F and numerous others. Assume the employer is no longer dealing with one opera singer, as in *Lumley v Guy*, but with large numbers of individuals performing similar work. Suppose the employer has a series of employment contracts with a group of miners – either fixed term or at will. If these employees acted collectively against the employer, they could seriously hold up the business for substantial sums. This is especially true in extractive industries, which historically have been, not by chance, the focal point of many labour disputes. Exit options for employers in extractive industries are often not viable. If one's coal is sitting there in the ground, one cannot simply take the business elsewhere. In effect, the business is a giant quasi-rent waiting to be collected by a rival entrepreneur of a special type, namely one who can enter into a contract, hold up the firm with a renegotiation threat, and then walk away with a lion's share of the take, even if it bankrupts the firm. To the extent that multiple individuals act cooperatively, the employer can no longer play off one worker against another. The employer is now dealing with a monopolist on the other side of the transaction, and thus finds itself in a situation fraught with peril.

These were precisely the circumstances in *Hitchman Coal & Coke v Mitchell*, which was decided in England in 1917, in which Justice Pitney, a most able if unappreciated Justice, applied the inducement of breach of contract theory in a labour dispute. A union knew that a number of men were working for an employer in a minefield under contracts at will, which meant that either party in each case could end the contract at any time without giving cause. It was also evident that an employee breached their contract if they joined a union but remained on the job, when the contract provided that the workers could not retain their jobs if they chose to join a union. These workers decided to remain in employment, but secretly agreed to go out on strike when the union called. The fig-leaf was that they merely promised to join the union at the key time, but were not members of it until they went out on strike.

The employer was apprised of this development, and could choose between two lines of attack. First, a private action for damages could be brought against every single employee. However, this would be enormously costly. Even if 'successful', the action would create such ill will in the workforce as to be self-defeating. The alternative was an action for inducement of breach of contract against the union. Here an employer can enjoin a single defendant, with which it generally had no ongoing relationship in any case, in order to protect its relationships with its workers, many of whom were skittish about resisting a union that they did not wish to join. By taking the inducement remedy, the employer reduces the transactions costs of legal action while increasing the power of the available remedy.

Justice Pitney considered the technical implications of the doctrine established by *Lumley v Guy*. He concluded that so long as the union had notice of the terms of the contract between employer and workers, and that these terms were perfectly open, the standard doctrines applied. Interestingly, many academic lawyers sympathetic to the union cause nonetheless agreed with the judge's textbook application of the tort of inducement of breach of contract in these circumstances.

From a more modern perspective, one of the attractions of the inducement action is that it serves a quasi-antitrust function. By granting the one injunctive remedy, it enables private enforcement against combinations and restraint of trade without the state needing to bring antitrust actions. One can avoid the difficult issue which dogged both American and English law throughout the nineteenth century – whether to treat a labour union as a criminal conspiracy, or at least a combination, in restraint of trade. The employer can in effect say: "You people can organise your union as something I cannot directly affect. But if so, it must be above board. I can now compete for workers by offering them a certain wage, perhaps involving a premium, on condition that they agree not to join a union". This is the so-called yellow dog contract.

Is it the kind of arrangement that could only result from unequal bargaining power? I very much doubt it. It is instructive that the employer in this situation did not have the power to offer a term contract, during which period the workers would not be able to join the union whether they quit their jobs or stayed. The final transaction was a straight at will relationship. If the workers all wanted to leave, join the union and take their chances in the bargaining game, they could immediately do so. But their chances of success were clearly diminished when they could no longer rely on coordinated action. And for that the external effects of the yellow dog contract are manifestly positive because it reduces the possibility of strike and dislocation in product markets.

Pro-union legislation: paradoxes of workplace democracy and the erosion of freedom

Understanding this paradigm puts us in a position to appreciate what happens if we attempt to go beyond the non-contextualised ABCs in labour disputes. Suppose one believes that a very different problem emerges as soon as one knows which party is the employer, which the employee, and which the union. In place of generic rules, the pressure rises to adopt contextual rules which take into account the peculiarities, needs and desires of each relevant party. There have been two different approaches along these lines – both erroneous and both destructive of the integrity of the market processes described so far.

The first is the English response. It did not come from any change of heart on the part of English common law judges regarding the tort of inducement of breach of contract. It was the Trade Disputes Act 1906 which suspended that tort insofar as it applied to employment contracts or labour disputes. A union could now take coercive collective action against an employer, and the employer had no means of fighting back. The Act may also have effectively prohibited binding contracts between individual workers and employers, to the extent that a unionisation question was left unresolved. This was an extremely powerful statute. For all practical purposes the ability to maintain competitive equilibrium in the labour market by private enforcement of actions was denied through class legislation.

The typical defence of such a statute is that it helps workers, who are needy and lack bargaining power. But that is true only in the very short run, and then only for certain privileged workers. The long-term consequences are adverse to workers as a class. The new legal regime influences the types of employment relationships into which employers are willing to enter – the activities they undertake, the location of their investment, the wages they offer, their desired mix of capital and labour, and many other factors. The standard maxim of economics applies to this case, as to so many others. If activist judges or legislators alter one term in a contractual arrangement, they cannot assume that other provisions – every other decision – will remain unchanged. Here, as elsewhere, firms respond to the incentives created by the new legal order.

In the case of employment contracts, legislators cannot expect employers to keep unaltered all the other contractual terms relating to employment, such as wages and conditions, or frequency of employment. If firms cannot fight, they will typically flee. Disinvestment will take place in certain industries. A statute designed to provide a transfer payment from one group of people to another, through changing common law rights and duties, in this case by disrupting contractual arrangements, ends up making both sides worse off than before.

Although developments in America were somewhat different, the results were broadly similar. Through the Norris-LaGuardia Act, passed in 1932 when Herbert Hoover was still president, America followed Britain in making a yellow dog contract no longer enforceable: injunctive relief was denied in labour disputes. A worker could therefore remain in the employ of a given employer while acquiring dual loyalties by joining a union. America did not follow Britain in simply stopping at that point, and allowing the fate of workplaces to be determined by economic muscle. The American framework was different. When Franklin Roosevelt took over from Hoover, his New Deal approach instituted a mandatory-bargaining relationship between union and employer in which the voice of the individual worker was effectively squeezed out.

The Wagner Act of 1935 introduced a regime of democratic voting into the workplace. Individual workers in a bargaining unit, as defined by law or by contract, could decide by supervised election whether they wanted a bargaining representative in the form of a union. If the workers in the bargaining unit rejected a union, the labour regime would be open bargaining. If they wanted a union, the selected union would take on the role of bargaining agent, not only for workers who had voted for the deal but also for workers who had opposed it. Moreover, under the early constructions of the statute, any private contract entered into by an individual employee and an employer that pre-dated union representation but which lasted beyond it was treated as inconsistent with the collective bargaining agreement, and therefore struck down as invalid.

The net effect of this ad hoc 'solution' to the tripartite relationship injected two new elements, both of which have created immense long-term difficulty. One element was democratic instability in industrial relations. There are great problems in achieving workable voting relationships in this strange new democracy, where all workers in a voting unit decide by majority rule whether to have a union. The first problem is specifying the coverage of a given union among a workforce. The gerrymandering of district boundaries in ordinary elections is a well-known problem. It also arises in the union context. Does one have a broad unit, at which point a union may have the support of only 45 percent of the 'electorate'? Or does one have a smaller unit, at which point the union could gain 60 percent support? Or does support for unionism increase with the size of the unit, so that the union will now opt for the larger unit while the employer pushes for the smaller? There are no satisfactory democratic rules determining eligibility to vote in elections of this type. It is not possible to have elections that decide who has the right to vote. Nor is there any clear principle of territoriality to sort out voters from non-voters. A huge administrative drag is imposed upon the labour market by this legal regime.

The second element concerns the conduct of these elections. We discover that once the principle of freedom of contract has been compromised, all other freedoms are necessarily compromised. Two examples illustrate this point.

Suppose an employer says to its workforce: "You are all free to join the union. But let me explain the pattern of disinvestment that will take place in this business in response to such activity". Despite the general guarantees concerning freedom of speech in the American Constitution, this speech turns out to be 'coercive' and thus is not protected. In other words, with a sufficiently eccentric definition of what people are allowed to do, a hothouse environment is created in which only rival worker and union factions can give their pitches to individual employees. The employer may be allowed to make certain promises but may not make certain types of threat. Nobody is quite sure where the baseline should be, since all the other common law baselines have been disrupted. Moreover, if union representation is decided under these circumstances, common law rules will necessarily be compromised as well.

Consider the rule that if somebody comes on to another person's land without the owner's consent, or remains after being asked to leave, that entrant is a trespasser. That rule of exclusive property is now weakened under the labour law. Certain types of union presence on employer premises are now protected by statute. This immediately puts the law on a collision course with the takings clause in the American Constitution, since one strong right of property owners is the ability to exclude other people from their property, except under cases of imminent peril, which are just not applicable here unless the success of a union is treated on a par with saving human life from flood or earthquakes.

I still remember the time when I first grasped this point. It was around 1962, when I was a junior at Columbia College taking a government course that dealt with constitutional issues. Our teacher, Hal Chase, had arranged for us to observe a morning's round of oral arguments before the Supreme Court. A lawyer representing an employer said: "Your honour, this case involves a common law trespass". I thought: "How sensible!". The late Justice Brennan looked at the lawyer and said: "So what?". To him the common law categories could not have mattered less. His casual remark helps us frame the relevant issues. An entire consistent network of contract and property relations is implicated in cases like *Lumley v Guy* that sport tripartite relationships. Once one alters the contractual arrangements in a collective bargaining negotiation, all the property rules and all the speech rules must change in sympathy with the initial alteration. The mixed states of the world are not stable. Hence if freedom of contract goes, inducement of breach of contract is allowed, freedom of speech becomes limited, and the right to exclude others must disappear, at least insofar as these relate to labour disputes. And just that type of systematic reworking took place in different ways in the American and English contexts.

What type of gains might justify such a radical change to the law? The usual explanation is very familiar: individual workers are helpless against employers, and need collective power in order to resist being exploited. Proponents of this view make the same mistake constantly – confusion of employer net worth with market structure. An individual employer may well possess vast wealth. But if so, it typically needs to hire, train and retain vast numbers of employees. If the employer is a monopolist, that can be handled by the old common law rules governing public utilities and common carriers: the employer is under an obligation to offer reasonable terms for the services demanded. But in employment relationships, there has never been a history of monopoly providers, because too many other parties can enter as employers and provide alternative outlets for labour. The railroad may have a monopoly with its passengers, but its clerks and drivers can work as well in other industries. The standard answer to the traditional concerns of labour unions is thus: "Do not worry about wealth. Worry about choice". If employers have free entry into a market, wages will be bid up without government intervention.

A profitable employer should be regarded, not as a threat, but as an opportunity. Their workers can generally have confidence they will be paid. Far from being exploited, in some ways those workers will be advantaged. The exact extent of this advantage is unclear. However, in unregulated markets, large firms sometimes do not need to pay the wage premiums required of small firms. Their workers have more confidence that their jobs will last beyond the next year or two, and do not demand the same implicit risk premium. This is not exploitation. It is simply rational risk assessment, the value of which should not be underrated. Thus the argument commonly made about inequality of bargaining power wholly misunderstands the employment relationship.

The perils of 'good faith' bargaining and the instability of monopolistic labour markets

There are other aspects to this problem. A strange inversion takes place whenever a union is created. To the extent that trade provides mutual benefits to the parties involved, it is competition which drives that process. Consequently, a sound legal framework should prevent the formation of combinations in restraint of trade. That was a standard English prohibition, ensuring that in various labour markets there would be choices on both sides. Yet early twentieth-century labour market policy goes in exactly the opposite direction – starting with a competitive labour market, until the law intervenes to create a monopolistic labour market. Labour and its supporters want a monopolistic market because then the legal system can regulate the bargaining behaviour of the two sides. Predictably enough, just as there are problems with union democracy, having created the bilateral monopoly between employer and employee, one needs to set up Roberts rules of orders, or Marquis rules of engagement in combat, to keep an appropriate balance between the two sides. After all, certain contractual terms cannot practicably be set by state fiat. One does not declare that all workers in all unions will earn exactly $5.78 per hour; that is infeasible. Instead one imposes an obligation to negotiate in good faith.

Not surprisingly, endowing the term 'good faith' with any genuine meaning turns out to be extremely difficult. One ends up instructing people: "Do not give take-it-or-leave-it offers. Do not have negotiations with low transactions costs. That is bad faith". In consequence, to negotiate the best contract within the law, one must conceal one's original preference – demand a huge sum, or offer a mere pittance, and then slowly yield to an intermediate position. Thus one way to demonstrate 'good faith' is to engage in acrimonious and lengthy contractual negotiations. Another is to make gradual concessions. These artificial and baroque negotiations are the direct, and unfortunate, outcome from this type of labour law.

A competitive market encourages a high rate of exchange with low transactions costs. In a monopolistic market there are high transactions costs and low rates of exchange. Yet somehow a monopolistic labour market is pronounced a social improvement on the basis of certain workers capturing short-term gains through monopoly rents, which will generally be dissipated over time as technological and investment patterns move against them.

The history of American labour law, and to some extent English, demonstrates that these monopolies do turn out to be unsustainable in the face of long-term substitution. In an open labour market, substitution is constantly taking place between one worker and another, or one contract and another. The moment substitution at that level is prohibited, dynamics are introduced similar to those involved in fixed exchange rates between countries. Many shocks that hit the system will not bring visible change. The marketplace will exhibit a certain rigidity. A union will push as hard as it can to extract benefits from an employer, with the employer using various tactics to fight back. But all of these monopoly arrangements are subject to cliff-like instabilities. Sooner or later organised labour will push too hard, and the business will simply decide to go elsewhere. Instead of making a large number of incremental adjustments, which is characteristic of competitive markets with multiple transactions, the firm experiences a few traumatic and discontinuous changes. Typically this means runaway factories and a major crisis.

Thus the second major problem with this type of labour regime is that it changes the nature of negotiations. Instead of frequent negotiations with small changes, as happens on any well-functioning exchange, one has a relatively small number of changes that are fairly dramatic, even traumatic, in their implications, leading to industrial instability. Once employers are aware of this problem, they attempt to expand their businesses in ways that minimise the difficulties. There is systematic substitution of capital for labour. Businesses move overseas in preference to hiring domestic labour. Age factionisation of the workforce takes place, to make solidarity between employees more difficult. The personalities who advance on the employer side gravitate to those whose unyielding toughness can endure the abuse of long-term hostile negotiations, as more reasonable types flee to other lines of work.

Moreover, the inefficiencies associated with unions are sufficiently damaging to workers themselves that the last 15 to 20 years have seen a rapid decline in union membership in virtually every advanced industrial country, except in the public sector. This is despite an essentially unchanged legal environment for unions since the 1930s or '40s, when many of these statutory protections were introduced.

The folly of mandatory terms

Developments in more recent times have reflected the failure of unionisation to capture the imagination of today's workers. Instead of institutionalised labour-market conflict, operating as a potential threat in all markets and realised only in some, a different attempt has been made to create monopolistic structures. The employment law no longer concerns itself with the complications of workers forming unions in their imagined extractive struggle with employers, but instead attempts to standardise contractual terms through government action. Thus an employer cannot hire individuals save on terms dictated by the government in respect of certain key elements of the contract. For instance, both New Zealand and the United States have age discrimination laws which prevent an employer from hiring somebody on a contract known to terminate at a given age. Both countries also have legislation concerning other aspects of the labour relationship, such as minimum wage laws and unjust dismissal laws.

It is characteristic of all these rules that they make no attempt to differentiate one firm from another. The state is now a monopoly arbiter of contractual terms. Since everybody must have the same term, there are important dimensions over which competition can no longer take place. This is not an argument that certain terms are in themselves undesirable, such as those allowing a worker to remain employed beyond age 65, or terms providing parental leave or various medical benefits. One of the reasons these elements are now so often mandated by statute or common law adjudication is that many progressive and enlightened employers do provide these benefits voluntarily. But the promoters of these laws have fallen into the usual fallacy of imagining that if one firm is willing to adopt a practice voluntarily, the arrangement must be efficient everywhere, and can therefore be mandated for all firms. These terms may well be appropriate for certain types of employment relationship, but for others they may be utterly inappropriate.

The composition of a workforce, and the demands of employers, are constantly changing. A competitive labour market involves incremental change, in which various contractual terms are tried and adopted in some segments of the market and rejected in others. These trends may converge to an equilibrium, or as new technology is introduced there might need to be a further series of adjustments. There is no reason for activist legislators to be involved. If they remember the eleventh commandment they will understand that these varying contractual arrangements are for the mutual benefit of their respective parties by virtue of their pedigree in voluntary consent. When the same terms are imposed upon other individuals, they violate that commandment because their source is government constraint rather than voluntary cooperation. They now operate as an implicit tax on the contractual relationship, and lead to dislocations similar to the collective bargaining tax.

Conclusion: Do your ABCs

These examples provide a very powerful lesson. Whether one is attempting to design a system of contract law, tort law or property law, the injunction is the same: go back to the old Oxford syllabus and do the ABCs. This is not teaching children how to read, but teaching lawyers how to think. Strangely enough, the older methods of generalisation, which ignored all differences between the parties, were the right way to approach the subject of contractual freedom. The simple explanation is that A, B and C all know who they are. If they have unique features, they express them through the contracts they make. Individuation takes place through voluntary assignment rather than state characterisation. One achieves finer differentiation, and greater contextualisation, because there may be two firms in the same industry which employ similar people but prefer somewhat different combinations of contractual terms.

When we understand that how firms write one term in a contract will depend heavily on how they are allowed to write the other terms, we recognise that interfering with one dimension of an employment relationship will adversely affect all the other dimensions. The employment contract should not be regarded as an exception, isolated from other types of contract. There should not be one type of contract for partners and another for employees. One must go back to the boring fundamentals, keep the law simple and coherent, and allow all the complexity to arise out of the actions of the people who know something about the transactions. Judges and legislators can know precious little about these complexities, and they do best when they do precious little.

Questions

Might the 1906 Trade Disputes Act have partially represented a belief that employers find it easier to collude anti-competitively than employees, who are greater in number? Often the rhetoric of the 'struggle' assumes that arrangements reached at the employers' 'club' about the going wage might interfere with proper market wage-setting just as much as union power.

There will need to be a very large number of people involved in wage-fixing at that club before effective cartelisation is feasible. Those employers who compete with one another for a given type of labour are generally not just confined to producers of a certain range of goods or services. A computer expert can work not only for a telecommunications company but for any employer who needs computer skills. In other words, the relevant market is much more broadly defined in employment than in product lines. Four Canadian potash producers who happen to occupy mines within five miles of one another may be able to collude in the potash market but they will not be able to collude in the labour market.

But suppose these concerns are justified, and collusion exists. My second point is that one cannot improve efficiency by imposing mandated terms on a monopoly. Even a monopolist is better off negotiating the ideal terms of contract. It will choose the right terms of engagement while changing only the wage term. It will not choose inefficient terms. If a government imposes inefficient terms on top of that, there is no improvement at all. The government has simply created a different distortion, giving us two problems – the monopoly problem and the mandatory term problem. There is no reason to believe that one distortion can be an antidote to the other. Indeed they will to some extent be additive. In the unlikely event of finding collusion by employers in labour markets, the correct response is an antitrust action against those employers.

Of course nobody today can bring an action against a union for collusion, because a union is always regarded as a privileged monopoly. I would prefer not even to rely on tort law or government action. If employers could write yellow dog contracts, such actions would not be needed. In effect a private antitrust attorney-general would achieve the desired result – a much more efficient solution than those in Western countries today. Wage levels would be higher. Employers would be prepared to pay more, owing to the greater stability in their expectations concerning investments in individual workers. Concerns about workers leaving could be addressed without the law being involved. Employers would simply back-load an employee's compensation or provide other inducements to stay longer. The experts on compensation packages already understand these solutions. They are the people to consult on this problem rather than turning to positive law.

I thus remain passionately non-interventionist on this issue. There are no parallels to problems such as occur in some network industries. These are essentially monopoly problems which cannot be escaped merely by breaking companies up. They pose genuine difficulties. But it is important to understand where the analysis is easy and where it is hard. The most contentious questions about labour relationships are the easiest to answer from a formal perspective. That is a great irony. Unless one believes that only workers compete in the marketplace, and not employers, one simply follows standard welfare economics and genuflects in the appropriate direction.

Where does the term 'yellow dog contract' come from?

I believe that it comes from a Jack London story. A yellow dog is a cowardly cur. A worker 'yellow dogs' when entering into an agreement with an employer not to join a union by supposedly sacrificing loyalty to fellow workers. It is a term like 'scab'. But I have used it so often that for me it is somewhat like the word 'Yank', which has lost its original derisive connotations. 'Yellow dog' is simply a convenient technical label to describe a certain set of contractual provisions. But in the American labour battles of the 1920s and 1930s, to be in favour of a yellow dog contract meant one was virtually wearing jackboots and giving a certain salute.

Back in 1983, I presented a paper at a conference at the Yale Law School celebrating the 50th anniversary of the New Deal. I told them that I would celebrate the New Deal by raining on the parade, and my paper defended the yellow dog contract. A law professor named Julius Getman declared that there was no difference between Richard Epstein and the Polish Armed Guards, who at that time were literally beating the workers of the Solidarity trade union into submission. I replied that in my ideal world no government employer would ever have monopoly power over a shipyard. One needs to analyse the full relationship in an employment contract rather than assume such violent propensities in me.

To the extent that the government is a sole employer, there is already a monopoly element on one side of the transaction. That is one of the real dangers of socialism. The great temptation is to conclude that fairness requires a bilateral monopoly rather than a single monopoly. But the employer then has nowhere to hide, because it is almost impossible to separate the economic from the political. One reason the size of the public sector should be minimised is to avoid confusing the role of government as sovereign with its role as employer. It is simply too difficult to disentangle the two. In the absence of complete separation, one often prefers an independent crown corporation – anything to create some distance.

Ironically, union power turns out to be greatest in the network industries, where they can take advantage of their monopoly position. No stockpiling is usually possible in these markets. One cannot take the train to work five times on Monday because the workers might go out on strike on Tuesday. Disruption is instantaneous and can be truly devastating. Some of us have argued that a union should never be permitted to shut down a public facility because of the negative externalities. We are told in response: "Life is tough for the workers. This action is important to them. The general public must simply learn to bear with the disruption". Meanwhile, factors that in the traditional competitive market were not externalities at all are promoted into such.

We thus see a surreal reversal of the normal world of third party effects. Genuine disruptions of contract are ignored; ordinary competitive pressures leading labour to move from one employer to another are now regarded as devastating externalities requiring the antidote of collective action. And as already mentioned, the moment legislators change one element in this mix, every single perception concerning the other elements also changes – what we mean by speech, by coercion, by democracy and by freedom. It is truly remarkable. One should never underestimate that Pickwickian development. Even if it is no part of anybody's intention, every contractual alteration eventually radiates. American labour law may be bad on these points. But at least it is internally consistent, and that is the problem.

The question as to whether the right to sue for personal injury should be restored brings out the relationship between contract and tort. There is a concern that the courts might simply override contracts voluntarily agreed by employer and employee if the right to sue is brought back. Have you any further thoughts on how we might progress this?

I have discussed the contract-tort dichotomy in connection with labour relations. Over this same period there is a parallel movement with respect to the allocation of risks associated with accident. The original view was that risk with respect to injury was an element in the contract of employment, and subject to bargaining just like wages or unionisation. Perhaps erroneously, the typical default position was that if an employer maintained their facilities negligently, they were obliged to pay in tort. However, if a worker was injured by a fellow employee who themselves had been negligent, the risk fell on the worker rather than the employer.

This is a more sensible risk allocation that is commonly imagined, provided workers are members of a union at the time or otherwise have say over with whom they work. One of the genuine functions of unions is that they organise team production amongst workers. That is one reason employers often want unions, within certain limits. If one wanted to hire a crew for a joint project, one would prefer an existing group of, for example, six people who had worked together for a long time rather than six individual workers who were thrown together for a single occasion. Moreover, if workers bear their own liability they will be very careful in selecting their workmates. This will have a powerful effect on accident levels.

The mode of extraction in mining industries in England strongly influenced the rules of contracting. Where individual workers went down a seam themselves, and there was relatively little team production, the rule was assumption of risk. With team production there was often social insurance in the form of a voluntary workmen's compensation plan. The common law employment rules were often criticised as demonstrating employer dictation. Yet this is not borne out by English practice in the last third of the nineteenth century. The two most dangerous industries were the railroads and the mines. Both industries saw voluntary workmen's compensation programmes introduced in order to socialise the risk in those cases where team production was important. This was employee-driven. That is inconceivable if one believes the domination story in which employers maximise welfare by minimising their liability. But it is perfectly plausible if the purpose of a contract is to maximise joint expected welfare. An employer can in effect guarantee the safety of the services they provide by giving a warranty for them – a form of damages payment. The wage can be correspondingly lowered in such a way that both sides are left better off. That is what happened.

Then regulation was introduced in England in the form of the Employers Liability Act 1880. One question was whether employment contracts entered into prior to the statute would survive under the new law. *Griffiths v Earl of Dudley*, decided in 1882, held that the contracts were good. But the 1897 workers' compensation statute explicitly banned these contracts. Voluntary arrangements between employment parties were forced to meet a set of statutory minimums that removed all opportunities for sensible bargaining. The United States at that time, and then Britain in 1910, introduced statutory workers' compensation systems that were modelled on the voluntary arrangements. They discarded tort law and assumption of risk, and kept it that way.

Unfortunately, systems introduced by statute are generally less stable than systems introduced by contract. One reason is the inability to control the coverage provisions: one runs into genuine problems with respect to cumulative trauma and occupational diseases. How are these conditions defined? How bad do they need to be? How much coverage should there be? What if the conditions are caused by multiple employers? What if they are caused by general atmospheric conditions? When legislators attempt to regulate these elements by statute, they typically run adrift. The second problem is that if benefit levels are wrong, all the incentives associated with the contracts will be wrong. Unless one is very careful, legislation will typically adopt the wrong benefit levels. It is another illustration of the dangers of the one-size-fits-all approach.

The third problem is that mandatory terms clearly have anti-competitive features. Two firms in an industry with different modes of production may have differently structured workers' compensation arrangements, and both of these might be efficient. If mandatory terms are imposed across that industry, there might be no cost to one firm but a tax on the other. Relative power tends to shift to larger firms which can amortise the cost of running a system over a broader base of employees.

In the United States, the whole contest now concerns whether freedom of contract prevails over accidental losses or personal injuries arising out of an employment relationship. The Federal Employer Liability Act, dating from the 1900s, prohibited contracting out of these requirements, and the Supreme Court upheld that restriction. Thus the erosion of freedom to contract was becoming part of the legal culture, even at a time when the courts were quite sympathetic to yellow dog contracts and distinguished between health issues, where regulation was permitted, and labour market issues such as the structure of the employment relationship, where it was not.

Today we see other evidences of the same problem. The New Zealand accident compensation system has its own share of problems including a huge unfunded liability, but it has one very good feature. New Zealand essentially had a tiny medical malpractice and product liability law. Properly understood, medical malpractice and product liability are contractual problems rather than tort problems. One typically wants a solution similar to that in workmen's compensation. Systematic under-compensation *ex post* will produce maximum gains to the parties *ex ante*. New Zealand is not far from the optimal solution. By contrast, American law is much more complicated and punitive, with almost no positive deterrent effect identified by anybody. New Zealand law does badly on road accidents, where a system based on negligence or strict liability is best. However, it retains stability in two markets, medical malpractice and product liability, where an assumption of risk defence should be paramount.

If New Zealand brings back a tort system, it is an open question whether a statute can be drafted that prevents those opposed to freedom of contract from creating legal havoc in the product liability and medical malpractice areas. The response of American businesses is to reject any programme that imposes any type of partial liability on them in, say, pharmaceutical cases, when their only defences are contractual. They do not believe that any contract ever drafted will stand up after a serious injury, when the problem of *ex post* regret becomes dominant. In consequence they insist upon, and sometimes obtain, statutory authorisation and protections of the contract, with an understanding that they will be compensated if liability is imposed in the teeth of these arrangements.

Abolishing the accident compensation scheme (ACC) in New Zealand would be a good move. However, a range of statutory waivers would need to be negotiated. A simple certification of freedom to contract will be overridden in the courts on all the grounds of imperfect information, inequality of bargaining power, standard form contracts, unconscionability, unfair surprise and incomplete disclosure that are always brought out by people who believe deep down that bargains are theft. Statutory waivers are the best means of handling the problem, but it remains a serious concern. By freezing tort law in New Zealand in a relatively good position, the accident compensation system brought gains that were far from trivial.

The real downside of the ACC in New Zealand has nothing to do with the incentive effects of tort law. It is a funding problem. We have the same problem in the United States: once government funding starts, those paying for a scheme lose the ability to control the benefits. It leads to a NZ$6 billion unfunded liability problem in New Zealand, and a US$400 billion, or more, problem in the United States in the form of programmes such as Medicare. Public funding in these areas typically begins with a resolution that a given scheme will be actuarially sound and solvent, yet always ends up abandoning that intention under political pressure. That is simply a fact of political life, and people should not fall for the same bland assurances twice. The first time might be regarded as a misfortune, but the second seems like carelessness or worse.

In a New Zealand context, contemporary contract law – especially labour law – seems to impose an idea of a mutuality in the employment relationship. Warm fuzzies such as good faith obligations are promoted as a contractual obligation, and a deterrent for each party to consider their interests in negotiating a contract. Does this not undermine the idea of an efficient labour market?

Mutuality has a long history in contract law. One example is the old rule of consideration: one party to a contract is not bound unless the other party reciprocally agrees to be bound. In other words, there must be mutuality of obligation for a contract to be formed. This doctrine confuses common practice in bargains with necessary practice. There can be a firm offer in which one party agrees to keep the offer open, and the other party is not obliged to accept it. That is in the interests of the offerer because the offerer expects a return, not because of an obligation on the other side to provide it. Over time people have learned that mutuality is a rule of thumb, but should not be an ultimate test of fairness. It must yield to an expression of contrary intention.

In the enforcement of remedies, the old rule used to be that if one side could obtain specific performance, so could the other. But sometimes the obligations are sufficiently asymmetrical as to make that rule nonsensical, and it has been jettisoned. When parties are silent about an arrangement, it is useful to consider whether mutuality or reciprocity is a test of fairness. In the example you raise, the good faith obligation does not arise from the contract. It is imposed from without, and that is dangerous. There are cases where the imposition of such an obligation is reasonable, but they have nothing to do with age discrimination legislation or other mandatory terms. The problem is rather sequential performance and opportunistic breach.

Suppose Mr A hires Ms B as a real estate saleswoman. Just as she is about to close a sale, Mr A fires her and completes the transaction himself. Is the money received by Mr A on this transaction, after firing Ms B, owed to her in equity? The good faith rule is designed to stop Mr A from expropriating her labour by a strategic firing mid-course. That has nothing to do with mandatory terms imposed *ex ante* upon individuals who would prefer different voluntary arrangements.

Thus I am not opposed to good faith obligations as a contractual supplement when the sequence of performance is important. That is its principal role. But when two people come together voluntarily, I cannot understand why good faith requires Mr A to employ Ms B after age 65, if he does not want that arrangement. There are no problems with interaction, strategic behaviour or opportunism, so it cannot be justified by the need for contractual stability. Thus good faith must be understood in terms of a specific contractual difficulty that it is designed to overcome. And typically when the drafters of real estate contracts see the possibility of a strategic firing as a recurring problem, they cease relying on the good faith rule. They adopt an explicit rule that in the event of Mr A firing Ms B before a transaction is completed, he pays her *quantum meruit* for the work done.

Legislators cannot be more explicit in defining good and bad faith. One cannot eliminate it, but for me good faith is essentially a default provision. Typically where one has knowledge of a specific field, one obtains more precise rules, such as schedules detailing how much should be paid, under what circumstances, and why. That process of individuation through contract reinforces the point made earlier. If the law begins with A, B and C, A and B can create a unique agreement if they have specific knowledge about the circumstances of their own transaction, including the associated risks.

You mentioned that damages are sometimes inadequate and cannot truly approximate the loss suffered. New Zealand and English courts seem reluctant to order specific performance in situations where a common understanding might have been expected. Is that the same in America and, if so, why?

Specific performance is frequently ordered in contracts involving the sale of real estate, for the obvious reasons. Supervision is very low, there is complete execution of the agreement, and clean-up damages will suffice. The courts are reluctant to order specific performance in cases where forcing an employer to keep an unwanted employee is regarded as a genuine affront to the principle of freedom to contract. That becomes the risk principle. Ironically there is one exception involving labour contracts. Specific performance never featured in traditional private labour law. An employer might have been able to obtain an injunction against an employee working for a different employer, but they could not obtain specific performance. The exception is union contracts. Now an employer must take workers back if a court finds an unfair labour practice or bad faith behaviour, as defined by the statute.

The rationale is clear enough. Workers are not regarded simply as workers, but as citizens in a given labour community. If an employer can fire them systematically, that employer can change the composition of the bargaining unit and its voting characteristics, thus helping it to defeat the union monopoly. Thus stability inside the labour democracy supposedly requires forced association of workers with respect to an employer. This illustrates once again our rule of inversion. As with speech, trespass and contractual freedom, so with remedies: negative externalities are fostered and positive externalities are curtailed.

With damages in employment relationships, the problem is typically calculating consequential damages. If an employee no longer comes to work, how are consequential damages determined? What if the employee would soon have left anyway? How should the training costs be treated? What is the genuine hypothetical state of affairs that would have held if the breach had never taken place? Such reconstruction is very difficult. Private agreements tend to have liquidated damage provisions, which are crude estimates rather than sophisticated attempts to answer such questions. But if we know that certain remedies are inefficient, it is pointless to rely on them in preference to an injunction against the third party, which reduces the likelihood of these remedies being exercised. The logic driving the common law rules appears sound.