

The Law of Contract Modifications: The Uncertain Quest for a Bench Mark of Enforceability

Varouj A. Aivazian

Michael J. Trebilcock

Michael Penny

Follow this and additional works at: <http://digitalcommons.osgoode.yorku.ca/ohlj>
Article

Citation Information

Aivazian, Varouj A.; Trebilcock, Michael J.; and Penny, Michael. "The Law of Contract Modifications: The Uncertain Quest for a Bench Mark of Enforceability." *Osgoode Hall Law Journal* 22.2 (1984) : 173-212.
<http://digitalcommons.osgoode.yorku.ca/ohlj/vol22/iss2/1>

This Article is brought to you for free and open access by the Journals at Osgoode Digital Commons. It has been accepted for inclusion in Osgoode Hall Law Journal by an authorized editor of Osgoode Digital Commons.

The Law of Contract Modifications: The Uncertain Quest for a Bench Mark of Enforceability

Abstract

The law of contract modifications poses an analytical paradox: Modifications should be presumptively invalid because they may encourage extortionary, coercive, opportunistic or monopolistic behaviour. Modifications should be presumptively valid because they represent the parties' assessment of their own best interests. A summary of the case-law reveals theoretical difficulties in the tests of enforceability. The authors use an economic framework of analysis in which they argue that the law of contract modification may be used to minimize the costs of contracting. Dynamic efficiency considerations argue against enforcement of all modifications although static efficiency considerations point in the opposite direction. The potential for opportunistic behaviour and "moral hazard" problems suggest some variables that are relevant to formulating legal rules which will aid in reducing the long-run costs of contracting.

THE LAW OF CONTRACT MODIFICATIONS: THE UNCERTAIN QUEST FOR A BENCH MARK OF ENFORCEABILITY

BY VAROUJ A. AIVAZIAN,* MICHAEL J. TREBILCOCK**
AND MICHAEL PENNY***

The law of contract modifications poses an analytical paradox: Modifications should be presumptively invalid because they may encourage extortionary, coercive, opportunistic or monopolistic behaviour. Modifications should be presumptively valid because they represent the parties' assessment of their own best interests. A summary of the case-law reveals theoretical difficulties in the tests of enforceability. The authors use an economic framework of analysis in which they argue that the law of contract modification may be used to minimize the costs of contracting. Dynamic efficiency considerations argue against enforcement of all modifications although static efficiency considerations point in the opposite direction. The potential for opportunistic behaviour and "moral hazard" problems suggest some variables that are relevant to formulating legal rules which will aid in reducing the long-run costs of contracting.

I. INTRODUCTION

The principles of law governing the enforceability of promises which modify pre-existing contractual obligations are currently undergoing what appears to be a dramatic transformation. Contractual modifications (mutually agreed changes in contractual terms made subsequent to the formation of the primary contract) have traditionally encountered problems of enforceability because of the requirement that enforceable promises must be given for consideration. Doctrine stipulates that a promise in exchange for a promise to perform an existing contractual obligation is, without more, unenforceable. This is often referred to as the pre-existing duty rule. In place of this rule, a growing

© Copyright 1984, V.A. Aivazian, M.J. Trebilcock, M. Penny.

* Associate Professor, Department of Economics and Faculty of Management Studies, University of Toronto.

** Professor and Director of the Law and Economics Programme, University of Toronto Law School.

*** B.A., L.L.B., Barrister and Solicitor, Supreme Court of Ontario.

This paper was initially presented at Professor J. S. Ziegel's 1983 Annual Commercial Law Workshop, held at the University of Toronto and subsequently at a Law and Economics Workshop at York University. Comments of workshop participants on earlier drafts are gratefully acknowledged, especially those of Professor Ian Macneil, the formal commentator at Professor Ziegel's Workshop. The comments of Professor Alan Schwartz are also gratefully acknowledged.

body of case law, statute law and legal commentary would enforce promises, particularly modification promises, without consideration, where they are "fair and equitable," made "voluntarily," not made in "bad faith" or as a result of "extortion" or "coercion," or are supported by "legitimate commercial reasons."

The nature of the apparent analytical paradox presented by contract modifications can be stated briefly: on the one hand, why would any party to a contract agree, by way of modification, to pay more or accept less than originally contracted for, without an appropriate *quid pro quo* (consideration), unless the other party had obtained bargaining power in the course of the relationship that he did not possess at the time of the contract formation and that he now seeks to exploit? If this explains most modification situations, then it might be argued that the law should attempt to discourage extortionary, coercive, opportunistic or monopolistic behaviour by refusing to enforce most modifications, perhaps by means of a presumption of invalidity. The traditional legal doctrine might be close to what is socially optimal. On the other hand, especially in commercial contexts where most litigated modification cases seem to arise, it might be argued that parties would typically not enter into modifications unless they both felt better off as a result relative to the position that would or might have been obtained without a modification. Hence, the law should respect the parties' assessment of what course of action best advances their joint welfare and enforce modifications, that is, apply a presumption of validity. This would reflect much contemporary legislative, judicial and academic thinking. Both propositions, despite being contradictory, have strong axiomatic appeal and hence, the apparent paradox that we attempt to deal with in this paper.

In wrestling with this apparent paradox the paper summarizes some of the older and more recent case-law that reflects the evolution in legal thinking on the enforceability of contract modifications and explores some of the theoretical difficulties which emerge from various attempts to articulate the content of the emerging enforceability tests. We cover briefly: the special case of satisfaction of debts; traditional doctrines of contract modification; the law of duress as it relates to the enforcement of contract modifications; and the weaknesses in recent formulations of the doctrine of economic duress in this context. We then attempt to develop and apply an economic framework of analysis to the major classes of situations involving contract modifications.

The economic framework of analysis that we develop takes as the objective of the legal rules governing contract modifications the minimization of the costs of contracting. In particular, we are concerned to

identify those legal rules that reduce transaction costs, both with respect to initial contract formation and subsequent contract modification surrounding contingencies that may affect the ultimate value of an exchange to the parties. The paradox described above is the product of a tension between two competing sets of efficiency considerations, which in some cases require difficult trade-offs that in part explain the ambivalence and confusion in the evolution of legal doctrine relating to contract modifications. Static efficiency considerations will generally require that contract modifications be enforced on the grounds that the immediate contracting parties perceive mutual gains from recontracting that cannot, at the time modification is proposed, be realized as fully by any alternative strategy. On the other hand, dynamic efficiency considerations focus on the long-run incentives for contracting parties at large imparted by a set of legal rules. In the modification context, these dynamic efficiency considerations adopt an *ex ante* perspective, rather than the *ex post* perspective implicit in the static efficiency considerations. Adopting the former perspective, rules that impose no constraints on recontracting may increase the over-all costs of contracting by creating incentives for opportunistic behaviour in cases where "hold-up" possibilities arise during contract performance. As well, even where a genuine change has occurred in the economic environment of the contract between the time of formation and the time of modification such that, in the absence of modification, one party faces an increase in the costs of performance relative to expectations at the time of contract performance, allowing recontracting may facilitate the reallocation of initially efficiently assigned risks. This leads to moral hazard problems that may attenuate incentives for efficient risk minimization or risk insurance strategies by the party who subsequently seeks the modification. Thus, what is in the best interests of two particular contracting parties *ex post* contract formation when a modification is proposed and what is in the interests *ex ante* of contracting parties generally in terms of legally ordained incentives and constraints that minimize the over-all costs of contracting may lead to divergent policy perspectives. Our framework of analysis emphasizes the dynamic or long-run incentive effects created by legal rules in the modification context and seeks to identify those rules that will reduce the long-run costs of contracting.

II. PAYMENT OF DEBTS

Since *Pinnel's Case*, a debtor's promise to pay part of a debt in return for the creditor's promise to forego the balance has been unenforceable because:

[P]ayment of a lesser sum on the day in satisfaction of a greater, cannot be any satisfaction for the whole, because it appears to the Judges that by no possibility a lesser sum can be a satisfaction to the plaintiff for a greater sum: but the gift of a horse, hawk or robe in satisfaction is good.¹

Pinnel's Case was applied in the leading decision of the House of Lords in *Foakes v. Beer*.² The utility of the rule was questioned even then by Lord Blackburn when he said:

What principally weighs me in thinking that Lord Coke made a mistake of fact is my conviction that all men of business, whether merchants or tradesman, do every day recognize and act on the ground that prompt payment of a part of their demand may be more beneficial to them than it would be to insist on their rights and enforce payment of the whole. Even where the debtor is perfectly solvent and sure to pay at last, this often is so. Where the credit of the debtor is doubtful it must be more so.³

Nevertheless, the rule in *Foakes v. Beer* has been followed in countless cases and was formally affirmed by the English Court of Appeal quite recently in *D & C Builders Ltd. v. Rees*.⁴ In that case, the plaintiff performed work for the defendant. The bill was not paid. The plaintiff became desperate for funds. The defendant, aware of the plaintiff's financial straits, offered to pay a substantially smaller sum, but only if it was expressly accepted in full satisfaction of the entire debt. The plaintiff accepted, received part payment and immediately informed the defendant that it did not consider the smaller payment "full satisfaction," relying, *inter alia*, on the rule in *Foakes v. Beer*. Lord Denning M. R. acceded to the formal authority of this rule but made it clear that the doctrine of equitable estoppel enunciated by him in *Central London Property Trust Ltd. v. Hightrees House Ltd.*⁵ operated as major exception to the strict rule. Lord Denning nevertheless denied the defendant the benefit of this principle because of the following important qualification: "The creditor is only barred from his legal rights when it would be *inequitable* for him to insist upon them . . . [H]e is not bound unless there has been truly an accord between them."⁶ Lord Denning found that the defendant had held the plaintiff to ransom, putting undue pressure on him: "She was making a threat to break the contract and she was doing it so as to compel the creditor to do what he was unwilling to do."⁷

¹ (1602), 5 Co. Rep. 117a, 77 E.R. 237.

² [1884] App. Cas. 605.

³ *Id.* at 622.

⁴ [1965] 3 All E.R. 837.

⁵ [1947] 1 K.B. 130.

⁶ *Supra* note 4, at 891.

⁷ *Id.*

In Ontario, the *Mercantile Law Amendment Act*⁸ statutorily reverses the rule in *Foakes v. Beer*. Section 16 provides:

Part performance of an obligation either before or after a breach thereof when expressly accepted by the creditor in satisfaction or rendered in pursuance of an agreement for that purpose though without any new consideration shall be held to extinguish the obligation.

Although it is unclear whether an "agreement" extracted by the exercise of undue pressure would qualify for enforcement under s. 16, it seems strongly arguable that any "agreement" extorted by conduct amounting to duress would be set aside.⁹

III. CONTRACT MODIFICATION

In contexts other than payment of debts the case usually cited as authority for the proposition that a promise in exchange for a promise to perform an existing contractual duty is unenforceable is *Stilk v. Myrick*.¹⁰ In *Stilk v. Myrick*, seamen were promised extra wages during a voyage and later brought an action to recover the extra amount promised. The Court was referred to the earlier case of *Harris v. Watson*¹¹ in which, on similar facts, recovery was refused. In *Harris*, Lord Kenyon said:

It has long been determined that when the freight is lost the wages are also lost. This rule was founded on a principle of policy, for if sailors were in all events to have their wages and in time of danger entitled to insist on an extra charge on such a promise as this, they would in many cases suffer a ship to sink, unless the captain would pay any extravagant demand they might think proper to make.¹²

In *Stilk*, Lord Ellenborough likewise refused the plaintiffs recovery of the extra amounts but based his findings on another principle of law:

I think *Harris v. Watson* was rightly decided; but I doubt whether the ground of public policy, upon which Lord Kenyon is stated to have proceeded, be the true principle on which the decision is to be supported. Here, I say, the agreement is void for want of consideration. . . . If they had been at liberty to quit the vessel at Cronstadt the case would have been quite different; or if the captain had capriciously discharged the two men who were wanting, the others might not have been compellable to take the whole duty upon themselves, and their agreeing to do so might have been a sufficient consideration for the promise.¹³

⁸ R.S.O. 1980, c.265.

⁹ See Waddams, *The Law of Contracts* (1977) at 92-93.

¹⁰ (1809), 2 Camp. 317, 170 E.R. 1168.

¹¹ (1791), Peake 102, 170 E.R. 94.

¹² *Id.* at 103 (Peake), 94 (E.R.).

¹³ *Supra* note 10, at 319 (Camp.), 1169 (E.R.); the report of Espinasse (1809), 6 Esp. 129, states that the principle of public policy in *Harris* was recognized as a just and proper policy;

Under existing modern doctrine, the requirement of fresh consideration for contract modifications is said to be subject to four formal exceptions:

- i) the doctrine of waiver;
- ii) the doctrine of promissory estoppel;
- iii) the *Mercantile Law Amendment Act*; and,
- iv) modification under seal.¹⁴

In addition, courts have employed various avoidance techniques to circumvent the nullifying effect of the consideration requirement and to enforce contract-modifying promises in some circumstances. Such avoidance techniques include:

- i) finding that the promisee himself has promised to do more than he was obligated to do under the existing contract, thus supplying the requisite consideration;
- ii) finding that the circumstances had changed so radically that either the contractual obligations are at an end (so that the second identical promise is consideration for a new contract) or that the second promise does, in such circumstances, constitute a promise to do more;
- iii) rather than modifying the original contract, the parties have rescinded it, and the "modifying" promise is actually part of a new contract.

In the recent Ontario case of *Gilbert Steel Ltd. v. University Construction Ltd.*,¹⁵ however, the traditional *Stilk v. Myrick* reasoning was applied to hold unenforceable the defendant's oral promise to pay a price for fabricated steel higher than that specified in the written contract. The Court of Appeal upheld the finding of the trial judge that there was no consideration for the defendant's oral promise. The Court further held that there had not been an implied rescission of the old contract and an assumption of new obligations under a new contract. The Court also rejected an argument that affording increased credit to the defendant was consideration for the increase in price. Finally, the Court rejected the plaintiff's claim that the defendant, by accepting delivery and making regular payments on account (albeit only totalling

doubts have been expressed as to the accuracy of Espinasse's report of *Stilk v. Myrick*, see Percy, (1981), 59 Can. B. Rev. 857 at 858.

¹⁴ Ontario Law Reform Commission, 1 *Report on Sale of Goods* (1979) at 97-98.

¹⁵ (1976), 12 O.R. (2d) 18 (C.A.); for a recent U.S. case which reached the same result on similar facts (but where a threat was made to suspend sales in the absence of a price increase), see *Roth Steel Products v. Sharon Steel Corporation*, 705 F. 2d 134 (6th Cir. 1983).

the original contract price), had acquiesced in the increase and was therefore estopped from asserting that it was not bound by the oral agreement. Two grounds were advanced for this conclusion: first, a plaintiff cannot found his cause of action in estoppel because estoppel acts as a shield and not a sword; second, the plaintiff had not discharged the burden of showing that there had been acquiescence and that the plaintiff had detrimentally relied on the defendant's apparent acceptance. Duress, of any kind, was not considered because of the trial judge's finding of fact that there had been no such duress.

This case has been much criticized in both result and reasoning.¹⁶ It is instructive to note that the reasons in *Gilbert Steel* have been questioned by the Ontario Divisional Court in *Re Tudale Explorations Ltd. v. Bruce*¹⁷ and not mentioned at all in the somewhat similar case of *Owen Sound Public Library Board v. Mail Developments Ltd.*¹⁸ It is also interesting to note the regret displayed by the trial judge in *Gilbert Steel* where, although conceding that the modification promise made commercial sense and that the defendant had acted in bad faith, he felt compelled to apply the consideration rule against the plaintiff. The general critique of the rule would maintain that the exalted position of consideration (now encrusted with exceptions and distorted by "avoidance techniques") ought to give way to a test of enforceability based on a clear articulation and application of the assumptions which underlie the traditional consideration requirement.¹⁹

One formulation of the general argument against the consideration requirement states that promises which modify existing and valid commercial contracts ought to be enforced as long as they are not extracted under duress or by unconscionable conduct.²⁰ Given the extensive and confusing array of exceptions to the consideration requirement, this argument has considerable persuasive force and intuitive appeal.

This reformulation of the rule of enforcement, however, while focusing attention on the right questions, is not complete until the content of a concept of duress (or any of the other frequently used labels) can be stated in a way which gives practical assistance to jurists, law-

¹⁶ Swan, "Consideration and the Reasons for Enforcing Contracts," in Reiter and Swan, eds., *Studies in Contract Law* (1980); Reiter, *Courts, Consideration and Common Sense* (1977), 27 U. Toronto L.J. 439.

¹⁷ (1978), 20 O.R. (2d) 593.

¹⁸ (1979), 26 O.R. 459 (C.A.).

¹⁹ Waddams, *supra* note 9; Swan, *supra* note 16; Reiter, *supra* note 16; Freid, *Contract as Promise* (1981) at 38-39; Eisenberg, *The Principles of Consideration* (1982), 67 Cornell L. Rev. 640.

²⁰ Waddams, *supra* note 9, at 90; Swan, *supra* note 16, at 29.

yers and contracting parties in furthering the goals of consistency and certainty in the law. Generalizations based on reasonableness or good faith do not supply the needed conceptual content. This problem is more specifically discussed, in the context of the emerging doctrine of economic duress, in the next two sections of this paper.

IV. ECONOMIC DURESS

The effect of duress on the enforceability of contracts evolved in the context of rigidly defined and strictly applied categories. These categories were:

- i) actual or threatened violence to the person — a contract entered into under duress of this kind was voidable;
- ii) improper application of the legal process — good faith suits and settlements were not included but transactions induced by the *improper* use of legal process would be set aside;
- iii) duress of goods — the wrongful seizure of property, followed by the demand for some payment or performance of some other act by the rightful owner as a condition of its return; and,
- iv) money paid pursuant to demands made by persons charged with the performance of public duties.

Pressure of an economic nature did not fall into a traditionally recognized category which would allow access to either a restitutionary or rescissionary remedy. A line of Australian cases, however, and recent English authorities²¹ adopting the reasoning of those cases, have expanded the traditional categories to include a concept of economic duress.²²

The leading Australian case is *Smith v. William Charlick Ltd.*²³ in which a payment was extracted from the plaintiff under threat of refusal to deal, a threat which, because of the defendant's monopoly over supply, would have put the plaintiff out of business. Although the plaintiff was refused recovery, Issacs J. did outline circumstances under which duress might be grounds for setting aside a contract:

"Compulsion" in relation to a payment of which refund is sought and whether it

²¹ These cases are helpfully reviewed by Ogilvie, *Wrongfulness, Rights and Economic Duress* (1984), 16 U. Ottawa L. Rev. 1.

²² For Canadian cases considering duress in contractual relations, see *Knutson v. Bourkes Syndicate*, [1941] S.C.R. 419; *Peter Kiewit v. Eakin Construction Limited*, [1960] S.C.R. 361; *George (Porky) Jacobs Enterprises Ltd. v. City of Regina*, [1964] S.C.R. 326; *Eadie v. The Corporation of the City of Brantford*, [1967] S.C.R. 573.

²³ (1924), 34 C.L.R. 38 (H.C.A.).

is also variously called "coercion", "extortion", "extraction", or "force", includes every species of duress or conduct analogous to duress, actual or threatened, exerted by or on behalf of the payee and applied to the person or the property or any right of the person who pays or, in some cases, of a person related to or in affinity with him.²⁴

The case of *Nixon v. Murphy*²⁵ further considered the phrase "duress . . . applied to the person or the property or any right of the person" in the context of a contract modification. The defendant, threatening contract repudiation, demanded from the plaintiff a sum of money in excess of the amount stipulated by an existing contract. The plaintiff paid the amount under protest and "without prejudice" in order to close the deal and meet its commitments to a third party. Long Innes J., holding that the plaintiff was entitled to be repaid the excess sum, said:

[T]he inferences of fact which I draw from this series of transactions are that the payment by the plaintiffs of this sum of money was induced by the defendant's threat to cancel the contract . . . that the plaintiffs did not make the payment voluntarily; nor intend to give up their right to the amount alleged by them not to be due nor to close the transaction; but that on the contrary the plaintiffs were endeavouring to the utmost of their ability to preserve their legal rights as regards the question in issue without jeopardizing their rights under the contract as against the defendants and without incurring further liability to their own sub-purchaser; and that the payment was made by the plaintiffs involuntarily and made under the compulsion of a threat of an unauthorized interference with the plaintiff's contractual and proprietary rights.²⁶

Recent English cases have confirmed and consolidated the Australian view of duress in relation to the enforceability of contracts. In *Occidental Worldwide Investment Corp. v. Stibs A/S Avanti, Stibs A/S Glarona, Stibs A/S Navalis (The "Siboen" and the "Sibotre")*²⁷ the plaintiff, to reduce the contract price for the charter of two tankers, falsely represented that it was on the brink of insolvency and that only a lower charter price would stave off liquidation. The defendant agreed to the modification but then subsequently withdrew the two vessels. The market rose and the defendant made healthy profits during the remainder of the original charter period. The plaintiff sued for damages based on the lower price. The defendant relied, *inter alia*, on the doctrine of duress. Kerr J. adopted the theory of the Australian line of cases, but held that although the defendant had agreed to the modifica-

²⁴ *Id.* at 56.

²⁵ (1925), 25 S.R. (N.S.W.) 151.

²⁶ *Id.* at 158. Both *Smith* and *Nixon* were followed in *T.A. Sundell & Sons Pty. Ltd. v. Emm Yannoulatos (Overseas) Pty. Ltd.*, [1956] S.R. (N.S.W.) 323. The Court in that case declined to restrict the ambit of the "duress" issue to the rigid categories outlined earlier in this paper, and quoted the passage from Issacs J. cited above.

²⁷ [1976] Lloyd's L.R. 293.

tion under great pressure, it was only commercial pressure, and not anything that could be said to coerce the will so as to vitiate consent. The plaintiff was nevertheless denied recovery because the contract modifications in its favour had been procured by fraudulent misrepresentations.

In *North Ocean Shipping Co. Ltd. v. Hyundai Construction Co. Ltd. (the "Atlantic Baron")*²⁸ the plaintiff contracted for the construction of a tanker. During construction, because of a devaluation in the American dollar, the defendant demanded a ten percent increase in the contract price. The plaintiff, requiring timely completion to fulfill its obligations under an advantageous charter contract with a third party, agreed to pay the increased amount "without prejudice." Mocatta J., again citing the Australian cases and recognizing that compulsion may take the form of "economic duress," held that the contract modification here had been procured by economic duress. A claim for damages, with all the uncertainties of litigation, was not a reasonable alternative to the plaintiff at the time, especially in light of the plaintiff's obligations to third parties. Nevertheless, recovery was denied because the plaintiff had failed to pursue its remedies promptly once the duress was no longer operative.

More recently, in *Pao On v. Lau Yin*,²⁹ the Privy Council considered, *inter alia*, the applicability of economic duress to a suit based on contractual provisions which were, arguably, modifications of a prior contract. This case concerned a large share transfer in Hong Kong. Although a contract was already in place, the plaintiff demanded a minimum price guarantee as a condition precedent to closing the deal. The defendants, although realizing the plaintiff could not successfully resist an action for specific performance, acceded to this demand in order to avoid the adverse market effect a large law suit would have on the shares. The defendants, however, subsequently refused to honour the guarantee and litigation resulted. The defence against enforcement of the guarantee was that it was extracted under duress. Lord Scarman recognized that economic pressure might render a contract voidable but held that such duress required a coercion of the will such as to vitiate consent. Commercial pressure alone was not enough.

In determining whether there was a coercion of will such that there was not true consent, it is material to enquire whether the person alleged to have been coerced did or did not protest; whether, at the time he was allegedly coerced into making

²⁸ [1978] 3 All E.R. 1170 (Q.B.D.); see Dadson, *Comment* (1980), *Fac. L. Rev. U. of T.* 223.

²⁹ [1979] 3 All E.R. 65 (P.C.).

the contract, he did or did not have an alternative course open to him such as an adequate legal remedy; whether he was independently advised; and whether after entering the contract, he took steps to avoid it.³⁰

Lord Scarman nevertheless made it clear that duress of an economic nature had become a recognized category of duress when he said:

[In] their Lordship's view, there is nothing contrary to principle in recognizing economic duress as a factor which may render a contract voidable, provided always that the basis of such recognition is that it must amount to a coercion of the will, which vitiates consent. It must be shown that the payment made on the contract entered into was not a voluntary act.³¹

The most recent English case of relevance to the emerging doctrine of economic duress is *Universe Tankships Inc. of Monrovia v. International Transport Workers Federation*.³² While not a case of contract modification, the reasons of Lord Diplock represent an interesting development in the explanation of economic duress and may be contrasted, analytically, with the reasons of Lord Scarman cited above. In *Universe Tankships* the appellant's ship was immobilized by the Federation's blacklisting of ships sailing under "flags of convenience." Although technically *obiter dictum*, Lord Diplock commented briefly on the issue of economic duress:

It is conceded that the financial consequences to the shipowners of the *Universe Sentinel* continuing to be rendered off hire under her time charter to Texaco, while the blacklisting continued, were so catastrophic as to amount to a coercion of the shipowners' will which vitiated their consents to those agreements and to the payments made by them to I.T.F. This concession makes it unnecessary for your Lordships to use the instant appeal as the occasion for a general consideration of the developing law of economic duress as a ground for treating contracts as voidable. . . . It is, however, crucial in my view to the decision of the instant appeal to identify the rationale of this development of the common law. It is not that the party seeking to avoid the contract which he has entered into with another party, or to recover money that he has paid to another party in response to a demand, did not know the nature or the precise terms of the contract at the time when he entered into it or did not understand the purpose for which the payment was demanded. The rationale is that his apparent consent was induced by pressure exercised upon him by that other party which the law does not regard as legitimate, with the consequence that the consent is treated in law as revocable unless approbated either expressly or by implication after the illegitimate pressure has ceased to operate on his mind.³³

³⁰ *Id.* at 78.

³¹ *Id.* at 79.

³² [1982] 2 W.L.R. 803 (H.L.).

³³ *Id.* at 812-13.

V. ANALYTICAL DIFFICULTIES IN MODERN THEORIES OF ECONOMIC DURESS

A. *The Unfettered Will*

In *Pao v. Lin Yau* "voluntary" consent is at the heart of determining whether illegitimate advantage has been taken by one party over another. Yet "volition" is a slippery, even metaphysical, concept. Devoid of readily ascertainable content, the same kinds of criticisms levied against the consideration requirement may easily surface against this approach to economic duress, that is, that the concept leads to conflicting decisions, insufficient predictability and anomalous results.

The basic theoretical problem in the *Pao On* account of duress is that it relies on a conceptually imprecise idea — that duress requires the will to be over-borne so that the resulting act is not a voluntary one.³⁴ Even in the "highwayman" example ("your money or your life") there is a sense in which it is perfectly correct to say that the victim *voluntarily* offers up his money as a result of a rational decision.³⁵ There is, in fact, at least one alternative. The victim of duress chooses to submit and fully intends to do so. As Dawson notes,³⁶ the more extreme the pressure, the more real the consent of the victim. Or as Dr. Samuel Johnson put the point somewhat more graphically, "Depend upon it, sir, when a man knows he is to be hanged in a fortnight, it concentrates his mind wonderfully."

B. *Illegitimate Pressure*

The language of Lord Diplock in *Universe Tankships* also raises problems of theoretical clarity when, after referring to the pressure of catastrophic financial consequences vitiating the consent of the victim, he goes on to refer to revocable consent induced by pressure "which the law does not regard as legitimate." This qualifying clause does nothing more than beg precisely the question that the inquiry must be directed to — how to formulate rational principles to distinguish legitimate from illegitimate pressure. The addition of a requirement that pressure must be legitimate before a modification will be enforced is an implicit recognition that the language of "consent" and "involuntariness", in

³⁴ For a full review of the inadequacy of this reasoning, see Atiyah, *Economic Duress and the "Overborne Will"* (1982), 98 L.Q.R. 197; for a critique and riposte, see, respectively, Tiplady, *Concept of Duress* (1983), 99 L.Q.R. 188 and Atiyah, *Duress and the Overborne Will Again* (1983), 99 L.Q.R. 353.

³⁵ See Kronman, *Contract Law and Distributive Justice* (1980), 89 Yale L.J. 472.

³⁶ *Economic Duress* (1947), 45 Mich. L. Rev. 253 at 267.

this context, lacks utility. But it is also a conclusory rather than an explicative qualification on the concept, and is itself unilluminating.

C. *Unconscionability*

Noted earlier were commentaries in which it has been argued persuasively that reliance upon the consideration requirement as a guide to determining which promises ought to be enforced has led to unsatisfactory and anomalous results. Consideration has been used as nothing more than a rough proxy for underlying assumptions about enforceable and non-enforceable promises, that is, whether or not they are "unfair" or "unconscionable" or "without legitimate commercial reason." Legal analysis, the argument goes, ought to go beyond this "proxy" approach and articulate explicitly those underlying assumptions which lead us to think that some promises ought to be enforced while others ought not to be, with attendant gains in consistency between cases, clarity of rationales and certainty of application.

Yet what is given by this analysis with one hand seems to be largely taken away with the other. Promises which modify existing contracts raise a particular concern in that, as time goes by, partial performance and increasing interdependence may change the parties' respective bargaining strengths. As a result, prior to complete performance, one of the contracting parties may engage in strategic or opportunistic behaviour in order to extort contractual concessions through the use of his newly acquired bargaining leverage. This is part one of the paradox outlined in the introduction to this paper. Why would any party agree to pay more or accept less unless he were the "victim" of such behaviour?

A common reason advanced for avoiding the enforcement of contract modifications is the presence of unconscionable behaviour - the exertion of economic duress or other forms of coercion which deprive the resulting promise of what is assumed to be its essential voluntary quality. For example, Treitel writes:

The rule that performance of an existing duty is no consideration sometimes serves the useful purpose of protecting one contracting party from extortion by the other. But in other cases the demand for extra pay may be perfectly reasonable, e.g. if a contractor who has agreed to do excavating work unexpectedly strikes hard rock. The fact that some promises of this kind may be obtained by extortion is no ground for invalidating them all. It would be better if the validity of these promises depended not on the doctrine of consideration but on the question whether the promisee had taken an unconscionable advantage of the promisor.³⁷

³⁷ Treitel, *The Law of Contracts* (5th ed., 1979) at 67.

The difficulty again is to explicate some content to the concept of "unconscionability" so that it becomes not only a tool of rationale discourse, but also a practical tool in the hands of judges and lawyers in determining which promises will or ought to be enforced.

The American approach to this problem is reflected in section 89 of the Second Restatement of the Law of Contracts (A.L.I., 1981) which states:

A promise modifying a duty under contract not fully performed on either side is binding —

- a) If the modification is fair and equitable in view of circumstances not anticipated by the parties when the contract was made; or
- b) to the extent provided by statute; or
- c) to the extent that justice requires enforcement in view of a material change of position in reliance on the promise.

The Comment explains that the requirement that a modification be fair and equitable goes beyond the absence of coercion and requires "an objectively demonstrable reason for seeking a modification." That reason must rest in circumstances not anticipated as part of the context in which the contract would be made and performed. Where such a reason exists, the relative financial strengths of the parties, the formalities attending the modification, and the extent to which its terms are performed or relied upon, may be relevant to determining whether the modification is binding.

Paragraph (b) refers principally to the Uniform Commercial Code, section 2-209, which applies the principle enunciated in the Restatement explicitly to sales contracts. Subsection 2-209(i) provides: "An agreement modifying a contract needs no consideration to be binding." The Official Comment on the section states:

[M]odifications made thereunder must meet the test of good faith imposed by this Act. The effective use of bad faith to escape performance of the original terms is barred, and the extortion of a 'modification' without legitimate commercial reason is ineffective as a violation of the duty of good faith. Nor can a mere technical consideration support a modification made in bad faith . . . [Good faith] may in some circumstances require an objectively demonstrable reason for seeking a modification . . . [A] market shift which makes performance come to involve a loss may provide such a reason even though there is no such unforeseen difficulty as would make out a legal excuse from performance.

The Ontario Law Reform Commission, in its recent Report on Sale of Goods, has recommended the adoption of a provision similar to section 2-209 of the U.C.C.³⁸

While the pre-U.C.C./Restatement law on modifications was

³⁸ *Supra* note 14, at 99ff.

much criticized for its restrictiveness, unrealism, and uncertainty as exceptions increasingly riddled the basic rule — “[the pre-existing duty rule] is, on the whole, that adjunct of the doctrine of consideration which has done most to give it a bad reputation”³⁹—much recent commentary is equally critical of judicial performance under the new rules, arguing that the case-law reflects no coherent theory of “good faith,” “extortion” or “coercion.” The result is decisions that are difficult to rationalize and rules that have become uncertain and unpredictable in their application.⁴⁰ While the American case law is amply canvassed in these commentaries, and is not reviewed here beyond drawing on some leading examples from the American case-law in Section VII of this paper, what is clear from the American experience is that the abolition of the consideration requirement and its attendant difficulties simply moves the inquiry to another difficult threshold. What is meant by “good faith,” “extortion” or “coercion” in this context? What is a “legitimate commercial reason” for modifying a contract?

VI. AN ECONOMIC FRAMEWORK OF ANALYSIS

In all contractual settings, parties' choices as to possible trading partners and/or terms of trade are subject to constraints. Concepts of “coercion” or “voluntariness”, standing alone, are not helpful in identifying the nature, source, and impact of that subset of constraints that render a resulting contract objectionable. In an attempt to break out of this vicious circle of inventing new language to restate old problems, we try, in the next two sections of this paper, to escape the sterility of metaphysical concepts, such as the “overborne will,” and the functional indeterminacy, in the present context, of recent, highly abstract, philosophical theorizing about the concept of “coercion”.⁴¹ We assume a rather prosaic objective for the law on modifications: minimizing transaction costs, and proceed to develop an economic framework that attempts to identify the variables that must be taken into account in formulating legal rules that advance this objective.

³⁹ Patterson, *An Apology for Consideration* (1958), 58 Colum. L. Rev. 929 at 936.

⁴⁰ See, e.g., Mather, *Contract Modification Under Duress* (1982), 33 So. Cal. L. Rev. 65; Robison, *Enforcing Extorted Contract Modifications* (1983), 68 Iowa L. Rev. 699; Hillman, *Policing Contract Modifications Under the U.C.C.* (1979), 64 Iowa L. Rev. 849; Brody, *Performance of a Pre-Existing Contractual Duty as Consideration* (1975), 52 Denver L.J. 433; Muris, *Opportunistic Behaviour and the Law of Contracts* (1981), 65 Minn. L. Rev. 521.

⁴¹ See Nozick, “Coercion,” in Morgenbesser *et al.*, eds., *Philosophy, Science and Method* (1969); Zimmerman, *Coercive Wage Offers* (1981), 10 Phil. Pub. Aff. 121; Alexander, *Zimmerman on Coercive Wage Offers* (1983), 12 Phil. Pub. Aff. 160; Zimmerman, *More on Coercive Wage Offers: A Reply to Alexander* (1983), 12 Phil. Pub. Aff. 165; Kronman, *supra* note 35.

We shall follow Posner⁴² and initially distinguish two alternative sets of cases in which contract modifications might be sought. In one set of cases there are no changes in the underlying economic conditions governing the initial contract except that the promisee has acquired some monopoly power *ex post* and exploits this power by forcing higher returns than provided for in the initial contract. Posner argues for the non-enforcement of contract modifications in such cases. The second set of cases is characterized by changes in the underlying economic conditions, or the emergence of new information about the underlying economic conditions governing the contract which prevent or inhibit the promisee from completing the promised performance without a modification of the contract. Posner argues that modification is justified in such cases because without that ability mutually advantageous exchanges may be precluded.

First, in relation to both sets of cases, it must be emphasized that the potential for opportunism⁴³ in the course of contractual performance is likely to be constrained in various ways. The party demanding a modification on threat of breach will need to take account of: the impact of this on future dealings with the other party if repeat transactions are envisaged;⁴⁴ the reputation effects on other potential trading partners in the market; ease of substitution by the party from whom the modification is demanded; initial contractual terms that may make the latter party unreceptive to a modification (for example, liquidated damage or penalty clauses, if enforceable, performance bonds, or back-end loading in payment schedules); the possibility of the latter party obtaining specific relief in the form of an injunction or specific performance; exposure to a damages claim in the event that modification is refused and breach occurs.⁴⁵ However, despite these constraints,

⁴² Posner, *Gratuitous Promises in Economics and Law* (1977), 6 J. Leg. Stud. 411.

⁴³ The use of the term opportunism may be misleading and is perhaps unfortunate. The term is not synonymous with self-interested behaviour. We assume self-interested behaviour throughout our discussion. However, such behaviour may or may not be opportunistic. By opportunistic behaviour we mean self-interested behaviour which is not appropriately constrained to result in *ex post* outcomes which are in the parties' mutual interest. By the same token, by non-opportunistic behaviour we mean self-interested behaviour appropriately constrained to result in collectively rational outcomes.

⁴⁴ See also Telser, *A Theory of Self-enforcing Agreements* (1980), 53(1) J. Bus. 27 for a discussion of factors that tend to make agreements self-enforcing, i.e., which tend to non-opportunistic behaviour. As Telser points out, at 27, "a self-enforcing agreement between two parties remains in force as long as each believes himself to be better off by continuing the agreement than he would by ending it." Strictly speaking, a contract cannot be self-enforcing since the existence of a contract implies third party intervention to determine whether a breach has occurred or to estimate damages from the breach. See also Williamson, *Credible Commitments: Using Hostages to Support Exchange* (1983), 73 Am. Econ. Rev. 519.

⁴⁵ For a discussion of how certain legal principles can deter opportunistic behaviour, see Muris, *Opportunistic Behaviour and the Law of Contracts* (1981), 65 U. Minn. L. Rev. 521.

there will be situations where there may be gains from engaging in opportunism — repeat transactions are not envisaged, market networks may imperfectly disseminate information about contractual performance, substitution may be difficult or costly, initial contractual provisions may not fully penalize or constrain opportunism and may be costly to negotiate in great detail, specific relief may be unavailable, and damages for breach may not fully compensate the non-breaching party for the costs associated with procuring a substitute and other consequential damages induced by the breach, or lack of exigible assets, limited liability or bankruptcy may preclude effective enforcement of a damages judgment.

Consider now the first set of cases. For these cases *ex post* contract modification is a zero sum game. What one party gains the other loses — in other words there is, by assumption, no room for co-operative recontracting. Assume an economic environment characterized by zero transaction costs, rational expectations and complete information about all contingencies. Assume also that the contract is initially drawn up under perfectly competitive conditions. We will argue that in such an environment whether the law enforces contract modifications is irrelevant from a resource allocation perspective and has no bearing on the economic welfare of the contracting parties. All that is required is that the law be *unambiguous*.

Suppose the law does enforce contract modifications. Then the terms of the initial contract will reflect the optimal future (*ex post*) strategy of the promisee. Suppose the promisor initially enters into the contract recognizing that the promisee's optimal future strategy will be to force higher returns through contract modifications. Since the law will enforce such modifications, the initial contract terms will be adjusted to reflect these future payoffs to the promisee. Potential extortionary monopoly rents will thus be fully impounded — given perfect competition and rational expectations — into *ex ante* contract terms.⁴⁶ As long as the promisee's optimal *ex post* strategy (whether opportunistic or not) is unique, it will be fully reflected in the initial contract in this rational expectations environment.

Suppose, on the other hand, that the law does *not* enforce contract modifications. Then the initial contractual terms will be adjusted so

⁴⁶ Thus, if an employee has the potential opportunity to demand higher wages on threat of breach, and finding a suitable replacement for the employee is expensive, then the employee's initial wage rate will adjust to reflect this "hold-up" potential. For discussion of the hold-up problem, see Klein, *Transaction Costs Determinants of "Unfair" Contractual Arrangements* (1980), 70 Am. Econ. Rev. 356.

that the return to the promisee, under the contract, will exceed his return in the previous case by a margin which exactly corresponds to the present value of the extortionary monopoly rents that would have accrued to the promisee if contract modifications were allowed and the promisee's optimal *ex post* strategy was to act opportunistically. However, if the promisee's optimal *ex post* strategy was to act non-opportunistically when the law allowed modifications, then obviously the return to the promisee will be the same in either legal setting. Hence, irrespective of whether the law enforces contract modifications, the return to the parties (in present value terms) will correspond to the same perfectly competitive allocation. As long as the law is unambiguous, resource allocation, as well as the economic welfare of the contracting parties, will be independent of whether or not the law allows contract modifications.⁴⁷

However, if there are positive transactions costs, then the choice of legal regime may have a significant influence on resource allocation. Initial analysis suggests that a law which unambiguously disallows contract modifications will generally be more efficient. Such a law economizes on the transaction costs that would otherwise be incurred in a regime that allowed contract modification. These costs not only include the direct costs associated with contract modifications (contract renegotiations take time and absorb other resources), but also those associated with various contractual and institutional arrangements designed to forestall opportunistic behaviour. These latter costs include those of writing, monitoring and enforcing detailed contractual provisions to penalize or constrain opportunism, as well as the costs of the resources expended by the promisee (for example bonding costs) to convince the promisor that he will not behave opportunistically.⁴⁸ It is in the promisee's interest to incur such costs (which limit his potential opportunism) in a regime that allows contract modifications to induce an optimal initial level of mutually advantageous contracts. However, such costs engender a reduction in the initial exchange opportunity since the exchange process is more costly (an inefficiency analogous to the "residual agency costs" discussed by Jensen and Meckling).⁴⁹ The par-

⁴⁷ The above discussion assumes that there are no unobservable "performance quality" asymmetries across promisees. With such asymmetries, an adverse selection problem may arise, as we argue below (see Akerlof, *The Market for "Lemons": Quality Uncertainty and the Market Mechanism* (1970), 84 Q. J. Econ. 488).

⁴⁸ Thus a promisee may invest in highly firm-specific capital which will have low salvage value if he cheats (produces low quality products) and loses future business. This maintains a firm's incentives to maintain quality and stay in business. See Clein and Leffler, *The Role of Market Forces Assuring Contractual Performance* (1981), 89 J. Pol. Econ. 615-41.

⁴⁹ For a discussion of the role of contractual provisions in deterring opportunistic behaviour

ties *ex ante* are therefore worse off than in the case where no modifications are allowed.

With unobservable "performance quality" asymmetries across promisees, an adverse selection problem⁵⁰ also arises in that promisees with non-opportunistic *ex post* strategies will not be adequately compensated for their above average performance "quality" and will withdraw from the exchange. Equilibrium contract terms will reward only those with opportunistic *ex post* strategies. A law which disallows contract modifications will be efficient since it will reduce the transactions costs associated with signalling or sorting activities designed to provide information on performance quality differences among promisees.

In short, in cases where contract modifications occur purely and simply as a result of changes in the strategic circumstances of the contracting parties, the enforcement or non-enforcement of modified contracts in a zero transaction costs environment with complete information about future contingencies and rational expectations will have no bearing on resource allocation or economic welfare. In an environment with positive transaction costs or incomplete information, a law which disallows contract modifications will economize on transaction costs and maximize the gains from contractual agreements. Hence, efficiency considerations dictate that contract modifications in this context be non-enforceable.

The second set of cases in which contract modifications may be sought (supervening changes in the economic environment of the contract) are those in which modifications can represent mutually advantageous positive-sum games. The Coase theorem⁵¹ implies that with zero transaction costs, Pareto efficient allocations — an allocation of resources is Pareto efficient if it is impossible to reallocate resources to make one or more individuals better off without making at least one other individual worse off — corresponding to such mutually advantageous (recontracting) exchanges will always emerge, provided that contract modifications are allowed by law. If recontracting between the parties in a particular *ex post* state of nature is mutually advantageous, then it will occur, leading to an optimal restructuring of contractual terms. These considerations suggest that contract modifications are necessary for the attainment of Pareto efficiency and should be allowed

on the part of a manager-agent and the nature of agency costs (monitoring, bonding, and residual costs) engendered by the agency relationship, see Jensen and Meckling, *Agency Costs and the Theory of the Firm*, [1976] 3 J. Fin. Econ. 305.

⁵⁰ See Akerlof, *supra* note 47.

⁵¹ See Coase, *The Problem of Social Cost*, [1960] 3 J. L. Econ. 1.

by law in this second set of cases. However, as we will see, such a conclusion is premature since there are additional considerations that bear on the problem.

Suppose that the law does *not* allow contract modifications. The impact of such a law on resource allocation and economic welfare depends very much on the economic environment postulated. Consider an economy characterized by complete contingent markets. In such an economy contracts can be written costlessly pre-specifying allocations among the contracting parties in every foreseeable contingency or state of nature, and contracts can be costlessly monitored and enforced by the parties.⁵² If modifications of contractual terms are disallowed, initially complete contingent contracts will provide for Pareto efficient allocations in every possible state of the world, precluding any need for *ex post* modifications of contractual terms.⁵³ Alternatively, in the absence of complete contingent contracts, legal rules, such as implied terms, damage remedies for breach, specific performance, or frustration, may be designed to induce Pareto optimal allocations in every *ex post* state of nature, again precluding any need for *ex post* modifications of contractual terms.⁵⁴

Once one abandons the assumption of complete contingent contracts,⁵⁵ or of legal remedies which fully substitute for contingent terms in contracts, the issue of whether the law should allow contract modification becomes more complicated. Consider an economy with incomplete contracts where *ex ante* contractual provisions do not exist to provide unambiguous Pareto optimal allocations in some states of nature. Suppose a promisee (supplier) of a product can avoid bankruptcy only if the promisor (buyer) agrees to a higher price than stipulated in the initial contract. If there are alternative suppliers of the product with identical or lower cost functions to the promisee, the non-enforceability of contract modifications will not affect resource allocation provided the transaction costs of turning to an alternative supplier are negligible. If

⁵² Such an economy is usually referred to as an "Arrow-Debreu" economy: see Arrow, *The Role of Securities in the Optimal Allocation of Risk Bearing* (1964), 31 Rev. Econ. Stud. 91 at 91-96, and Debreu, *The Theory of Value* (1959).

⁵³ Thus, *ex ante* contractual provisions can serve to attenuate in a complete contracts environment the misallocation of resources that may otherwise result from the non-enforceability of contract modifications.

⁵⁴ For a discussion of how legal remedies for breach of contract substitute for contingent terms in contracts, see Shavell, *The Design of Contracts and Remedies for Breach* (1984), 99 Q. J. Econ. 121.

⁵⁵ One should not readily dismiss contingent contracts as unrealistic since we do observe such contracts in some settings, e.g., cost-plus contracts. Their existence will depend on a trade-off between risk sharing benefits from contingent provisions versus moral hazard costs due to imperfect monitoring and verification of *ex post* states.

these transaction costs are non-negligible and if a suit for damages for breach will not fully compensate the non-breaching party for the costs associated with procuring a substitute, then non-enforceability of contract modifications may adversely affect resource allocation.⁵⁶ If alternative suppliers of the product have higher costs than the promisee, then contract modifications may be Pareto efficient. Of course, in this circumstance the modified price demanded by the promisee cannot exceed what his higher-cost competitors would charge, otherwise the promisor would turn to them. Thus, the smaller the promisee's effective monopoly power (reflected in the degree of availability of substitutes for his product and the effectiveness of a legal action for breach), the lower the scope for Pareto efficient contract modifications.

Another set of issues that arises once one abandons the complete markets environment relates to the question of efficient allocation of risk among the contracting parties in environments characterized by positive transaction costs and asymmetric distribution of information about future contingencies. A strong economic argument can be made that when risk assignments are not explicit (and third-party insurance possibilities are missing), as in incomplete markets, the law should assign the risk of an adverse outcome to the party who is the superior risk bearer.⁵⁷ Superior risk bearer simply means the party who can minimize the risk at least cost or, where risk is the product of exogenous events beyond the control of either party, can insure at least cost. Relative insurance efficiencies turn on relative ability (a) to appraise both the probability of a risk materializing and the magnitude of the costs that will follow from it, and (b) to spread the risk through market insurance or diversification.

Where the contract has specifically assigned certain risks, the presumption should be that they have been assigned to the superior (that is, most efficient) risk bearer and, moreover, that he has been adequately compensated for bearing them, thus removing any distributional objections to leaving him to bear the costs. Where the contract is not explicit about the allocation of a given risk, it seems a reasonable presumptive rule of interpretation to assume that the parties intended for it to be borne by the superior risk bearer and that it has simply been impounded in the categorical terms of the contract. In either case, the party bearing such risks should not be permitted to subsequently

⁵⁶ Also, if bankruptcy costs are positive, then the modification of the original contract serves to keep the promisee-firm in the industry, economizing on the deadweight administrative costs to society associated with bankruptcy.

⁵⁷ See Posner and Rosenfield, *Impossibility and Related Doctrines in Contract Law: An Economic Analysis* (1977), 6 J. Leg. Stud. 83.

reallocate them to the other party by contract modification, taking advantage of limitations in the relief available to the party on breach in order to induce the modification.

If the law allows contract modifications, it imposes at least part of the risks of adverse outcomes on the promisor. To the extent that the promisee can modify risks by his activities, and there is imperfect monitoring of his activities by the promisor, a moral hazard problem will arise. The promisee will have incentives to increase his risk exposure (relative to the case where he fully bears the risks) over time by devoting fewer resources to risk prevention than is optimal and increasing the probability of occurrence of adverse states (for example, bankruptcy) in which contract modifications may be necessary. Thus, the enforcement of contract modifications may generate inappropriate incentives over time that affect the probabilities of alternative outcomes or states of nature. Inappropriate incentives are created not only for the immediate contracting parties but, even more importantly, for future contracting parties in similar circumstances. These dynamic moral hazard inefficiencies, as well as the transaction costs associated with monitoring, bonding, and enforcement activity through a variety of contractual or institutional arrangements designed to overcome these inefficiencies,⁶⁸ have to be weighed against the static efficiencies resulting from contract modifications discussed earlier.

As noted, Posner, in his analysis of the modification cases, suggests a basic two-fold distinction between pure strategic modifications and those that are a response to changed circumstances rendering it uneconomic for one party to complete contractual performance as initially envisaged. Posner would not enforce modifications in the first class of case, but would in the second. The first he views as a species of monopoly where the object of the law should be to deter strategic behaviour; the second he views as excluding strategic considerations by virtue of the changes (or perceived changes) in underlying circumstances.

As noted, Posner's treatment of this second class of case, while perhaps superficially attractive, ignores a counterpart problem to that of strategic behaviour in the first class of cases, namely the problem of moral hazard; to allow risk reallocation through modification away from the superior risk bearer attenuates incentives to take efficient risk reduction or risk insurance precautions.

In developing our analysis of the category two modification cases,

⁶⁸ For example, the promisor may undertake costly monitoring of the promisee's activities to reduce potential risk shifting or the promisee may expend resources (e.g., bonding costs) to reduce his incentives for risk shifting.

we are, perhaps ironically, adapting in large part the framework of analysis developed by Posner and Rosenfield in the context of the doctrine of frustration/impossibility.⁵⁹ In effect, we view enforceable modifications as a substitute for the doctrine of frustration: contracting parties, facing the occurrence of some intervening event that substantially affects the cost of performance may, under some circumstances, rearrange their contractual rights and obligations either through invocation of the assistance of the courts pursuant to the doctrine of frustration or through private recontracting. In other cases, while the underlying factual circumstances of the contract may not have changed, new information about those circumstances may have been uncovered. The parties may have contracted on the basis of incomplete or inaccurate information about the underlying factual environment of the contract. This situation is the domain of the doctrine of mistake. The economic considerations bearing on permissible rearrangements by virtue of contract modifications, the doctrine of frustration, or the doctrine of mistake⁶⁰ would seem to be similar.

Under the doctrine of frustration, for example, the courts typically will not allow the doctrine to be invoked where the risk in question has been clearly assumed by one of the parties. There also appears to be some evidence that even where the risk is not clearly assigned, judicial decisions are broadly consistent with the view that the doctrine should not be invoked to reallocate risks away from the party who is the superior risk bearer as determined by objective circumstances.⁶¹

What is clear from this is that just as Posner would not permit frustration to occur in all cases of supervening events rendering contractual performance to one of the parties unprofitable, so it would be no more efficient to allow contractual modifications to reallocate risks in all such cases.

However, in applying this approach both to frustration and modification cases, several problems must be acknowledged. First, the contract may not clearly assign given risks and an objective inquiry into who is the superior risk bearer may sometimes be indeterminate (for example, one party may have superior ability to appraise the probability of a particular event occurring, but the other party may have superior ability to appraise the magnitude of the costs entailed if it does). Such a case might arise, for example, where there is a contract

⁵⁹ Posner and Rosenfield, *supra* note 57.

⁶⁰ See Kronman, *Mistake, Disclosure, Information, and the Law of Contracts* (1978), 7 J. Leg. Stud. 1.

⁶¹ Posner and Rosenfield, *supra* note 57.

to produce a machine to the buyer's specifications and a strike prevents timely completion. The supplier can probably best judge the likelihood of a strike in his industry and may to some extent be able to control or provide for this risk, while the buyer may best be able to judge the costs to him of delayed delivery. Here, even though the risk may not be an uncommon one, in the absence of a fully specified contract it may be difficult to know where the presumption leads that risks should be treated as initially assigned to the superior risk bearer and that subsequent reassignments through, for example, frustration or modification, should be foreclosed. However, one desirable effect of a no-modification rule in the particular example given may be to lead the machine manufacturer to negotiate a limited liquidated damages clause governing breach, thus leading to an efficient initial allocation of risks.

Another problem may arise in cases where it may be feasible to identify clearly the superior risk bearer but the risk in question is very remote. This is referred to in American law as the "unforeseen circumstances" exception to the traditional consideration requirement for valid modifications. The law in many contexts rather facilely tends to classify risks as foreseeable or unforeseeable, while economists would tend simply to view all risks as carrying some degree of probability with them. The fact that a risk is one of low probability does not necessarily mean that it has not been foreseen and appropriately discounted in the initial contract terms, or that there are not efficient risk reduction or risk insurance strategies available to the parties, or that in this latter respect one party is not a superior risk bearer to the other. Indeed, most insurance contracts involve low probability risks. However, it remains true that if the risk is remote enough, that is, carries a very low probability, the effects of assigning it to one party or the other may have little effect on contractual behaviour because the *expected* (that is, *ex ante*) costs to whoever bears the risk are so small as to warrant very little, if anything, in the way of efficient precautionary responses.

Taking these two problematic cases together — (a) indeterminacy in the identification of the superior risk bearer, and (b) highly remote risks — it can be argued that these are the transactional domains where contractual modifications should be permitted. In both cases, permitting modifications enables the party to whom a modification is proposed to capture some of the static gains from recontracting by avoiding losses he may well sustain in the event of a breach. Losses in dynamic efficiency associated with a rule that permits subsequent reassignment of initially efficiently assigned risks are likely to be small, in the first case because we cannot determine with confidence what is an efficient allocation of risks, and in the second case because while this

can be determined, disturbing the efficient allocation of risks will have little impact on the long-run behaviour of contracting parties at large while enabling the static gains for the immediate parties from recontracting, which may be significant, to be realized.

Permitting modifications in these two classes of cases, in effect, permits a flexible form of risk sharing by the contracting partners. Assuming the parties to be risk averse and the risk in question one that neither can readily control or insure against, contract modification may reduce the variance in possible contractual outcomes for the parties and thus be Pareto efficient.⁶² A countervailing consideration is, of course, that the promisor now is exposed to the risk of strategic behaviour by the promisee designed to exploit or manipulate these two exceptions and to the adjudication costs attendant on courts attempting to apply rules that do not supply knife-edged sharpness in the characterization of situations as falling within or outside given rules or exceptions thereto.

VII. APPLICATIONS OF THE ECONOMIC APPROACH

Following the economic framework of analysis outlined in the previous section of the paper, we categorize modification cases into two broad classes; (i) those involving pure strategic behaviour where there has been no change in the underlying economic environment causing either contracting party to revise his expectations of the gains from contractual performance relative to the expectations that were held when the contract was entered into; (ii) those where there has been a change in the underlying economic environment (or at least in the information available to the parties about the environment) causing at least one of the contracting parties to revise his expectations of the gains from contractual performance relative to the expectations held at the time of contract formation. We subdivide this second class of cases into situations where it seems efficient to allow the parties, through contract modification, to rearrange the contractual allocation of risks and those where it is not.

A. *Pure Strategic Modification: The Problem of Opportunism*

In this class of case, the economic concern with treating modifications as enforceable is that this creates incentives for opportunistic behaviour on the part of promisees; this in turn is economically wasteful because of the gaming (that is, transaction costs) likely to be generated

⁶² Cf. Posner and Rosenfield, *id.* Section III A.

in the recontracting process. In short, modifications entered into in these circumstances are not Pareto efficient forms of recontracting, but are at best zero-sum games. In addition, distributional or ethical objections may be raised to the redivision of the gains from trade as a result of opportunism by one party.

While, as we have earlier pointed out, there will often be constraints on such behaviour, in some cases they will not be fully binding. On breach there may be a significant risk that the non-breaching party cannot secure the relief required to put himself in the same position as if the contract had been fully performed. The other party may threaten breach as the inducement for accepting a modification that will prove less costly for the non-breaching party. However, even here the party threatening breach, given our assumption that there has been no change in the underlying economic environment that warrants any revision of the parties' initial expectations of the respective gains from trade, will typically also have to contemplate some loss from breach relative to contractual performance. Apart from reputation effects, he may face the risk of foregoing all or some of the profits of the initial transaction, so that the non-breaching party can in turn threaten to reject a proposed modification and invite a breach. In this event, the outcome would be a negative sum game — both parties would be worse off relative both to performance under the initial contract and to possible outcomes under agreed modifications to the initial contract. Yet, in the gaming environment that a law rendering contractual modifications enforceable in this context would create, this possibility cannot be foreclosed.

In the pure strategic modification context, a law rendering modifications unenforceable allows the potential opportunist two options: performance or breach, and performing in this context will almost always be more profitable than breaching (given our definition of pure strategic modifications). Thus, an optimal law would seem to be one that takes this form, promoting efficiency by eliminating incentives that lead to socially wasteful gaming behaviour and related transaction costs.

Because of the constraints we have noted, cases of pure strategic modifications are likely to be relatively rare. *Occidental Worldwide Investment Corp. v. Skibs A/S Avanti*,⁶³ involving an attempt to induce a reduction in charter-party rates by misrepresentation as to the charterer's financial circumstances, would seem to be one. Another example is the American case of *Austin Instrument Inc. v. Loral Corp.*⁶⁴ Here,

⁶³ See text accompanying note 27, *supra*.

⁶⁴ 29 N.Y. 2d 124, 272 N.E. 2d 533 (1971).

Austin had subcontracted with Loral to provide gear components to Loral for radar sets that the latter had contracted to supply to the U.S. Navy. This contract contained liquidated damage and cancellation clauses that were triggered if Loral fell behind its agreed schedule of deliveries. Loral was subsequently awarded a second contract with the Navy and Austin submitted a sub-contract for the gear components. Austin then threatened to breach its first sub-contract by suspending deliveries unless it received additional payments on this contract and unless it was awarded the second sub-contract. Loral acceded to these demands, after canvassing alternative suppliers without success. The New York Courts of Appeals allowed an action by Loral to recover extra payments on the first sub-contract on the grounds of duress. No alternative source of supply was available and conventional remedies for breach of contract were probably inadequate given the nature of Loral's exposure under the primary contract with the Navy. The case was viewed by a majority of the court as a pure "hold-up" situation where nothing had changed in the underlying circumstances of the sub-contract that could possibly justify the sub-contractor's demands. A further well-known American example of this first category of modification cases is *Alaska Packers Association v. Domenico*,⁶⁵ which concerned an agreement by sailors to perform services in navigating and operating a vessel in Alaskan waters. After reaching Alaska the sailors refused to perform their contracted services unless the defendant's representative signed an agreement to pay the plaintiff's additional wages. Being unable to procure other men because of the remoteness of the location, the defendant's supervisor complied. The Court found the second agreement void for want of consideration. The lower court had held that the defendant had waived its right to damages for breach of contract by entering into another contract for performance at a higher rate. The Court of Appeals held that there was no voluntary waiver on the part of the appellant of the breach of the original contract.

However, as other early maritime cases⁶⁶ tend to suggest, modification cases will rarely be this straightforward. In several of these cases, the sailor's claim to the higher wages that had been agreed to by way of contractual modification rested in part on a change of circumstances (for example, desertion, death, or illness of other crew-members that increased the work-load of, or at least the risks faced by, the remaining members of the crew). This makes such cases, at least potentially, category two cases where there has been a change in the under-

⁶⁵ 117 F. 99 (9th Circ. 1902).

⁶⁶ See text accompanying notes 10 and 11, *supra*.

lying economic environment since contract formation, and the general question is whether one party should be able to induce a reallocation of the risks of contractual performance through modification.

Our impression from a reading of both Anglo-Canadian and American case-law on contract modification is that in fact most cases are of the category two variety — in a sense not surprisingly because here one of the contracting parties not only has an incentive to exploit the gains from opportunism inherent in the other party's inability to secure full relief on breach but also faces a situation where only a contract modification is likely to yield him all or most of the gains from trade that he (perhaps optimistically) initially expected. In other words, his position, given performance of the initial contract, is substantially worse than that of a party in a category one case, relative to expected contractual returns. Moreover, the promisor, in the event of attempted enforcement of the initial contract, faces a real risk of breach or even bankruptcy by the promisee which may similarly create major divergences between the promisor's *ex ante* and *ex post* contractual returns. Thus, gains from modification in category two cases may be higher for both parties. We devote the balance of this section to more extended analysis of these cases.

B. *Modifications as Responses to Changes in the Economic Environment of a Contract: The Problem of Moral Hazard*

In this class of case, the analysis that we developed in the preceding section of the paper argues against the enforceability of modifications where this would permit an inefficient reallocation of initially efficiently assigned risks, and thus the attenuation of *ex ante* incentives to efficient risk reduction or insurance (the moral hazard problem). The promisee may be able to extract such modifications because of the weakness of informal constraints on his behaviour (for example, attenuated reputation effects, no expectation of repeat transactions) and of formal constraints, in particular inadequate remedies for breach. However, where the superior initial risk bearer cannot be confidently identified, or where he can be but the risk in question is very remote, modifications may be Pareto efficient in that losses to both parties associated with the breach can be at least partially avoided and long-run incentives for efficient risk-bearing by future contracting parties in the class of situation in question may not be significantly distorted. We now examine several leading cases or case-types in the light of this analytical framework.

One situation that has arisen quite frequently in the cases involves construction contracts. One type of case we have already dealt with —

where the builder "holds up" the site owner part way through performance, demanding extra payments by trading on the fact that the site owner will not be able to find ready substitutes and may possess incomplete remedies for breach. This is a category one case (pure strategic modifications) and, as indirectly illustrated by *Austin Instrument Inc. v. Loral Corp.*,⁶⁷ should not be enforced.

Another type of construction case involves situations where the builder, in the course of contractual performance, encounters, for example, difficult and costly excavation problems because of unusual soil conditions. In the case of a large builder erecting many commercial buildings, it seems obvious that he will typically be the superior risk bearer relative to the site owner. He can appraise the risks *ex ante* more efficiently and can often diversify them away across a number of similar projects. Some excavation projects will be more difficult, some less difficult, than the norm but these are familiar risks to him and he can be expected to adjust his contractual terms accordingly. Modifications that reassign the risks to the site owner attenuate the incentive for builders in this situation to act efficiently and should not be enforceable.⁶⁸ To enforce modifications in this context is tantamount to turning a fixed-price contract into a cost-plus contract, an arrangement the parties could have negotiated but manifestly did not, presumably for good reason given the desired allocation of risks.

On the other hand, in the case where a small contractor agrees, for a modest sum, to dig a cellar under the other party's house and encounters a hard crust of earth three feet deep under which is a quagmire of wet mud,⁶⁹ it is arguable that both qualifications to the superior risk bearer approach may apply. Determining who is the superior risk bearer may be difficult; first, these conditions may fall entirely outside the realm of experience of such contractors (making it difficult to know whether he or the home-owner could best appraise the risks or take risk reduction precautions), and second, the builder's ability to diversify across similar projects may be highly circumscribed by virtue of the size or nature of his operations. Even if he is clearly the superior risk bearer, to deny the possibility of a modification here imposes costs not only on him but on the homeowner who must search out a substitute and attempt to recover costs for breach in a damage claim which, even

⁶⁷ *Supra* note 64.

⁶⁸ See *King v. Duluth, Missabe and Northern Rly Co.*, 61 Minn. 482, 63 N.W. 1105 (1895), holding such a modification to be unenforceable.

⁶⁹ See *Linz v. Schuck*, 106 Md. 220, 67 A. 286 (1907), holding such a modification to be enforceable.

if successful, may yield less than full compensation (for example, in the event of the contractor's bankruptcy). By allowing the mutual gains from modification relative to breach to be realized, the question arises whether this significantly attenuates the incentives of contractors in this type of situation to take efficient risk reduction or risk insurance precautions. If the risk in question was highly remote and the *expected* costs associated with bearing the risk so small that no significant changes in behaviour are likely to be induced by shifting it, then permitting a modification seems to enhance the welfare of contracting parties and is efficient.

It should be added that in the above example an efficient modification is likely to be forthcoming whatever the homeowner's remedies on breach. Suppose the contractor agreed to undertake the job for \$2,000 and that the homeowner placed a value on the improvement of \$3,000 (a consumer surplus of \$1,000), but that in light of the discovery of the unusual soil conditions a contractor would require a competitive price of \$30,000 to undertake such a job. If the owner can compel specific performance, he will impose costs of \$30,000 on the builder but still only acquire an improvement which is worth \$3,000 to him. Any modification releasing the builder from his obligations that involves him paying the homeowner some amount above \$3,000 and less than \$30,000 will leave both parties better off. If the homeowner is confined to an action in damages for his expected losses, he can presumably only recover \$3,000, which leaves him indifferent between performance and non-performance, and an offer of a payment of this amount by the builder should induce a release of the latter's obligations. In both cases, the work originally contracted for will not be performed (an efficient outcome, when costs, as re-evaluated, are related to benefits), although the availability of particular remedies will affect the distributional impact of the modification on the parties. On the other hand, if the homeowner's consumer surplus is substantially in excess of the original contract price, a modification that provides for the work to proceed on payment of additional consideration to the builder, which together with the original price is less than the consumer surplus, is also efficient, subject to the assumption that the risk in question was very remote or the superior risk bearer difficult to identify (or both) and the dynamic inefficiency effects of risk reallocation by modification thus relatively inconsequential.

Other common types of modifications in the construction context involve changes by the site owner in the initial contract specifications during contractual performance. Suppose a homeowner has contracted for major renovations or extensions to his home, and as the work takes

shape decides that he would prefer larger windows than those originally specified. He is clearly the superior risk bearer in relation to the risk of a change in his own preferences and should not be able to use a contract modification to shift the risk to the builder. However, if the modification entails his bearing the full costs of the extra work, he has borne the costs associated with the occurrence of this risk and the modification is Pareto efficient — both parties are better off (or at least no worse off) as a result of the modification and enforcing the modification creates no long-run incentives for risk shifting by other contracting parties in similar situations. Thus, receipt of fresh consideration for a modification is important as an assurance that initially efficiently assigned risks have not been subsequently inefficiently reassigned.

This role of consideration as a test of efficient modifications may, however, raise some problematic evidentiary issues. Suppose, in the above example, that the builder claims that the homeowner inadequately compensated him for the new work, threatening delay in payment and exploiting the builder's cash flow problems to induce a highly favourable modification (by way of analogy with the facts in *D & C Builders v. Rees*).⁷⁰ If this claim is true, then the homeowner, through the modification, has succeeded in partially reallocating the risk of a change in his preferences to the builder. The converse situation might also arise. The builder demands an excessive sum for the new work, threatening to complete the work in accordance with the original specifications and compelling the owner to adopt much more costly substitute arrangements to modify the completed work (illustrating that opportunism is not restricted solely to category one situations). In each case, a modification on competitive terms is efficient but the actual modification that results may not be, in the first case because the owner has partially reallocated an initially efficiently assigned risk, in the second because transaction costs are generated by the builder's opportunities to game the homeowner and the latter's desire to avoid this.

In maritime salvage cases, the courts will often imply a term of reasonable compensation to avoid the exploitation of situational monopolies,⁷¹ although how easy such a term would be to operationalize in a modification context would seem to be largely a function of how easily competitive reference points can be established for purposes of comparison. This may not be so difficult in the window/construction example.

⁷⁰ *Supra* note 4.

⁷¹ See Trebilcock, "An Economic Approach to the Doctrine of Unconscionability," in Reiter and Swan, eds., *Studies in Contract Law* (1980) at 393, 394 and cases cited at footnote 54 therein.

However, in the example of a large-scale contract to develop and supply a new armament system to the government, where the latter changes its specifications during the contract for agreed additional consideration, it may be very difficult for a court subsequently to adjudicate claims by either side that their mutual interdependence created gaming possibilities that rendered the consideration either excessive or insufficient.

While it will be obvious from these examples that the presence or absence of consideration may be relevant to a determination of whether a modification is efficient, no simple rule can capture all possible factual nuances. Moreover, to invite or require the courts to scrutinize carefully the appropriateness of any new consideration payable on a modification may, in some cases, generate transaction (adjudication) costs that exceed those involved in unconstrained (and sometimes inefficient) modifications. For example, in the armament case, it might well be argued that the prospect of repeat dealings and reputation effects in a highly specialized and concentrated market sufficiently constrain incentives for opportunism by either party so that a strong presumption of enforceability should attach to an agreed modification. In more discrete transactions, such as that between the homeowner and the builder, closer scrutiny of the terms of agreed modifications may be warranted.

Various other types of long-term supply contracts for goods and services involving modifications induced by supervening events raise similar issues. Some leading examples of long-term contracts for the supply of goods which raise these issues are briefly reviewed.

In the American case of *Goebel v. Linn*⁷² the defendant brewers contracted for the supply of ice at a fixed rate. The ice was a crucial component of the defendant's business, being used to preserve the large beer stocks kept on hand. Due to a mild winter, the plaintiff could not accumulate all required stocks of ice and informed the defendant that it would not perform its supply contract unless the brewery agreed to pay a higher rate. The defendants agreed and took delivery of its ice but later repudiated the contract as modified. The Court held the contract modification enforceable, stating that the brewing company "chose for reasons which they must have deemed sufficient at the time to submit to the company's demand and pay the increased price rather than rely upon their strict rights under the existing contract." The Court went on to observe:

⁷² 47 Mich. 489, 11 N.W. 284 (1882).

Suppose, for example, the defendants had satisfied themselves that the ice company under the very extraordinary circumstances of the entire failure of the local crop of ice must be ruined if their existing contracts were to be insisted upon, and must be utterly unable to respond in damages, it is plain that then, whether they chose to rely upon their contract or not, it could have been of little or no value to them. Unexpected and extraordinary circumstances had rendered the contract worthless, and they must either make a new arrangement, or, in insisting on holding the ice company to the existing contract, they would ruin the ice company and thereby at the same time ruin themselves.⁷³

Posner endorses the result in this case on the basis of his distinction between modifications entered into where there has been a change in underlying circumstances and those where there has not. However, our emphasis on efficient risk allocation suggests a more cautious approach to evaluating the correctness of this decision. If mild winters were one of the occupational hazards of running an ice business during the era in question, the ice company would seem clearly to be the superior risk bearer, both in terms of risk reduction (for example, making different inventory or stand-by sub-contractual arrangements) and in terms of risk insurance (that is, being better able to appraise the impact of climatic variations on the supply of ice and adjusting the initial contractual fixed price accordingly). Only where the winter in question was quite out of the ordinary so that the gains from recontracting were likely to outweigh the long-run costs of moral hazard problems associated with permitting recontracting in these circumstances, could one support the decision. Otherwise, permitting recontracting is inefficient.

Another leading American decision which seems suspect for similar reasons is *Pirrone v. Monarch Wine Co. of Georgia*.⁷⁴ Pirrone, a modest-quantity wine-maker had agreed to sell to Monarch, America's largest purchaser of peach brandy, 150,000 gallons in 1968 and again in 1969. As it had over-bought brandy and because the market price fell to two-thirds of the contract price, Monarch breached its shipping schedule. In 1969, with Pirrone complaining of financial difficulties, its facilities loaded with brandy, its production of more profitable wine curtailed, and its ability to ship controlled by Monarch which was required to initiate a shipping permit, the parties agreed to a modification that provided for shipments to Monarch of the residue of the 1968 year but terminated Pirrone's rights in respect of the 1969 sale. The Court of Appeal for the Fifth Circuit refused an action by Pirrone for loss of profits on the 1969 sale, holding the termination agreement to be binding and dismissing a claim of duress.

⁷³ *Id.* at 285 (N.W.).

⁷⁴ 497 F. 2d 25 (5th Circ. 1974).

Clearly this seems to be a case where the risk in question was more efficiently borne by Monarch — the risk of a fall in the American market price for peach brandy (of which they were America's largest purchaser) was a risk that they could much better evaluate than Pirrone and at least partially diversify against. Moreover, the fluctuation in price seems not to have been entirely out of the ordinary and might reasonably be assumed to have been impounded in the terms of the original contract. The modification, in effect, reallocated this initially efficiently assigned risk to Pirrone, the less efficient risk bearer. This is a clear case where the enforcement of a modification generates significant moral hazard type incentives.

A more problematic case is *Gilbert Steel*,⁷⁵ discussed above, where it will be recalled that a price increase by way of modification to a steel supply contract was held unenforceable. The Court found no evidence of threats by the supplier to breach by suspending deliveries, but rather a desire on the part of the buyer to accede to the proposed price increase in order to obtain favourable treatment from the supplier under a subsequent contract. Moreover, there appears to have been some evidence of industry practice to the effect that price increases of this kind would normally be passed along, at least in part. Given the relational nature of the contract where the prospect of repeat transactions was likely to have constrained opportunistic behaviour on both sides, and given evidence of industry practice on cost pass-throughs, making judgments on the initially efficient allocation of this risk uncertain, it would seem arguable that the modification was Pareto efficient. The buyer's subsequent attempt to resile from it, having presumably derived the contemplated benefits from it, may itself be a form of opportunism — a "final period" problem in a long-term relationship, where one party decides to terminate the relationship and acts opportunistically towards the other party who is proceeding on the converse assumption. Game theory literature suggests that this problem and associated problems of "backward induction" relate principally to long-term relationships with a known terminal date, as opposed to long-term relationships whose terminal date is uncertain where theory suggests that there will be strong incentives for co-operative outcomes to emerge.⁷⁶

While the relational nature of the contract in *Gilbert Steel* may thus provide a mild immunizing presumption towards the modification there in issue, that the distinction between discrete and relational con-

⁷⁵ See text accompanying note 15, *supra*.

⁷⁶ See Luce and Raiffa, *Games and Decisions* (1952); Axelrod, *The Evolution of Co-operation* (1984).

tracts cannot be elevated into a sharp doctrinal divide is demonstrated both by game theory and by cases such as *Austin*⁷⁷ and *Pirrone*⁷⁸ where the modifications seemed clearly inefficient. Long-term relationships often ultimately come to an end, and not always harmoniously. Changes in such a relationship of the kind exemplified in proposed contract modifications may be part of the termination process. Indeed, modifications are unlikely to be the subject of litigation unless termination has occurred. Whether such modifications, if accepted, should be enforced will then have to be judged against the back-drop of an aborted and hostile relationship. Moreover, as Macneil has pointed out, any simple definitional dichotomy between discrete and relational contracts is impossible to frame.⁷⁹ Degrees of integration and interdependence in commercial relationships reflect enormous variations.

Two other well-known cases involving modifications to long-term contracts further illustrate the framework of analysis developed in this paper. In the first, *Central London Property Trust Ltd. v. High Trees House Ltd.*,⁸⁰ a ninety-nine year lease of an apartment block entered into in 1937 was modified in 1940 by a reduction in rents when the lessee found himself unable to let many of the apartments on account of war-time conditions. This modification was held enforceable as long as these conditions prevailed. Similarly, in *Raggow v. Scougall and Co.*,⁸¹ the plaintiff entered into a five year employment contract for a fixed salary as the defendant's designer. The contract was entered into in 1913, World War I broke out in 1914 and many of the defendant's customers cancelled their orders, compelling the defendants to consider closing their business altogether. Instead, they agreed with most of their employees, including the plaintiff, on wage reductions for so long as the war continued. Despite this agreement, the plaintiff subsequently sought to enforce his original contract at the higher salary. The action was rejected.

The risk in question in each case — world war — was presumably very remote or highly unusual. It seems equally unlikely that this risk had been explicitly or implicitly allocated by the contract or that one party was the superior risk bearer of it than the other. Thus, the modification in each case could scarcely be said to have reallocated initially

⁷⁷ See text accompanying note 64, *supra*.

⁷⁸ See text accompanying note 73, *supra*.

⁷⁹ Macneil, *Economic Analysis of Contractual Relations* (1981), 75 N.W.U.L. Rev. 1018, esp. at 1025 *et seq.*

⁸⁰ [1947] K.B. 130.

⁸¹ (1915), 31 T.L.R. 564.

efficiently assigned risks, while it enabled *both* parties to avoid losses that probably would have been sustained if performance of the initial contract had been sought and breach occurred. Thus, the modifications were Pareto efficient for the parties in a short-run sense and appear to have created no undesirable long-run incentives for inefficient risk reallocation.

Another type of case that is not wholly dissimilar in some respects to *Raggow v. Scougall* presents some problematic features for us, and warrants an extended comment. A enters into a long-term contract to employ B (for example, a baseball player) at a fixed wage. In the course of his contract, B receives an offer of employment from C to work at a higher wage. B proposes modification of his contract with A to offset in whole or in part C's higher wage offer. Here, in contrast to the situation in *Raggow v. Scougall*, B's services are worth more in the contingency that has materialized, not less. Should such a modification, if accepted, be enforced?⁸²

Posner argues that modifications in this type of example should be enforceable because C values B's services more highly than A did under the original contract and thus it would be efficient to allow breach, subject to damages.⁸³ B should be entitled to compensation from A for foregoing this opportunity. We are not convinced of the logic or relevance of this argument.

In our analysis, the case seems clearly to involve a category two situation. The cost to B of performance under the contract with A has risen since contract formation because his opportunity costs have risen since contract formation (treating a benefit foregone as economically similar to a cost incurred). The question would then arise as to whether the contract had clearly assigned this risk, in this case an upward fluctuation in the market wage rate for B's services, to B or if not, whether B could be independently identified as the superior risk bearer. It might be argued that B is the superior risk bearer because he knows better than anyone else how he is likely to perform in the future and can control the factors influencing his performance better than anyone else. On the other hand, the team owner in the baseball example may be better able to assess the impact on team performance and spectator appeal of individual performance fluctuations or be better able to diversify these risks away across players. If B is clearly the superior risk bearer, the question would still arise as to whether the risk in question

⁸² See *Schwartzreich v. Baumann - Basch Inc.* 231 N.Y. 196, 131 N.E. 887 (1921), holding such a modification to be enforceable.

⁸³ *Supra* note 42.

is a very remote risk, the shifting of which to A by way of modification will not significantly attenuate B's incentives to efficient risk reduction precautions, in this case by searching out market opportunities before contracting with A, not afterwards, and maintaining consistent levels of performance thereafter. In the professional sports case, better third party offers would not seem to be extraordinary events and would generally seem to be risks that have been assigned in long-term, fixed wage contracts.

In the case of long-term sporting contracts at a fixed wage, there are, of course, symmetric risks — the player may play better than expected, in which event his contract wage will be below the market wage, or he may play worse, in which event his contract wage will be above his market wage. While it may be in the *ex ante* interests of both employer and employee to insulate themselves from the risk of upward and downward fluctuations in the market wage rate for a player, the effectiveness of remedies for breach may be asymmetric. If an owner attempts to unload a slumping player or reduce his wage, he will subject himself to a damage claim that will probably capture most of the player's losses and can probably be effectively enforced. If the player seeks to breach his contract and take up a more lucrative offer with another team, the first team owner may have difficulty proving various forms of consequential losses (for example, reduced gate receipts as a result of the player's departure), effectively enforcing his judgment if B's principal assets are human capital, or obtaining specific performance. It is these limitations on A's remedies for breach that gives B some bargaining leverage with respect to a contract modification.

If the matter rested there, framing optimal modification rules would be a highly problematic exercise. Freely allowing modifications would effectively convert a long-term fixed wage contract into a short-term variable wage contract, at least on one side and probably eventually on both. But, by assumption this is what neither side initially wanted. On the other hand, a rule of non-enforcement may leave the owner with the worst of all possible worlds — no effective remedy for breach if B threatens to leave, and no ability to "bribe" B to induce him to stay.

Fortunately, in this type of situation, the law has been sensitive to these considerations and is likely to make available to A a negative injunction preventing B from taking up C's offer while not compelling him to play for A.⁸⁴ This sharply reduces B's bargaining leverage vis-à-

⁸⁴ See, e.g., *Warner Bros. v. Nelson*, [1937] 1 K.B. 209.

vis A in demanding a modification and forces C to buy-out the injunction by compensating A, which he will presumably only do if B is worth more to C than A (thus ensuring only “efficient” breaches). Thus, in these circumstances, if B is clearly the superior risk bearer and the risk in question is not highly remote, a policy of non-enforcement of contractual modification may be efficient in a long-run sense if these contracts are not to unravel into short-term variable wage contracts which neither side regards as optimal. However, little may turn on the matter if a form of specific relief is available to the owner. Such modifications as are agreed to are unlikely to affect major reallocations of risk — the owner need not accept them. It is possible, of course, that even where B is subject to a negative injunction, he may be able to threaten to play poorly for A (and, for example, demoralize the team), which may be a difficult contractual breach to prove or police. Thus, a “sweetener” may be mutually advantageous, irrespective of any remaining divergence between the modified contract and market wage rates. The important point remains that to the extent that A’s remedies for breach approach completeness, the choice of rule on the enforceability of contractual modifications becomes largely inconsequential, although a presumption of mutual short and long-term gains from modification becomes more plausible in this scenario. Paradoxically, to the extent that A’s remedies for breach are *radically* incomplete, modification also becomes a correspondingly more valuable short-run option to him (as well as to B), despite the fact that in the long-run the A’s and B’s of this world may all end up worse off — a classic “prisoner’s dilemma” problem that some constraints on modification may properly attempt to alleviate. These constraints will always be difficult to enforce in this context because of the incentives to immediate contracting parties to attempt to substitute around them by alternative institutional arrangements, whatever the long-run costs to contracting parties at large from permitting inefficient reallocations of risks.

The last class of case that we examine is the debtor-creditor case of part payment, which has attracted divergent legal rules over time, as we have earlier noted. Our framework suggests the following analysis. These cases may fall into category one or category two. Suppose the following situation: Debtor (D) owes creditor (C) \$200. D offers to pay C \$100 in full discharge. C accepts, knowing that, for small debts, litigation costs and difficulties of enforcing judgment generate an expected return from action for breach of \$90. D trades on this fact in offering \$100 in full settlement. This is a pure strategic modification case and, under our analysis, the modification should not be enforceable. If this were the rule, D now faces a choice between performing (paying \$200)

and facing litigation where if any of the transaction costs involved in suit are borne by him, his exposure will be greater than \$200 (even though C's net return is only \$90). In most cases, it will be rational for him to perform.

On the other hand, suppose D, because of intervening illness or unemployment, cannot presently afford to pay more than \$100. Should a modification to this effect be enforceable? As to who is the superior risk bearer in relation to these contingencies, it might be argued that they are risks over which debtors typically have little control and that creditors have some information in aggregate on such risks and are better able to diversify them away across similar transactions. This would suggest a case for enforceability. Moreover, if there was empirical evidence that by far the greatest proportion of personal debtor delinquencies fell into this category, there might perhaps be an economic case for a single rule (such as that in the *Mercantile Law Amendment Act*)⁸⁵ so as to economize on transaction costs involved in courts discriminating between situations on a case-by-case basis.

VIII. CONCLUSIONS

Our analysis assumes that in the pure strategic modification cases the static efficiency gains to the immediate contracting parties from modification relative to breach are outweighed by the long-run or dynamic efficiency losses from encouraging opportunistic behaviour. Moreover, in most pure strategic modification cases a rule rendering modifications unenforceable will ensure contractual performance rather than breach - the economically optimal state of affairs. In the changed circumstances cases, the trade-off between static and dynamic efficiency is not so straightforward. Precluding modification in all such cases runs the real risk of breach (and attendant costs), given our assumption here of changes in underlying cost conditions (or perceptions thereof). However, to allow modification in all such cases, while permitting the parties to realize the static efficiency gains from recontracting relative to breach, creates long-run or dynamic efficiency losses as a result of the attenuation of incentives to efficient risk reduction or insurance as well as generating transaction costs on recontracting. Only in cases where the efficient allocation of risks is indeterminate, both subjectively and objectively, or where the risk in question is extremely remote so that the expected costs of bearing it do not induce significant efficient precautionary responses, is it likely that the static efficiency

⁸⁵ See text accompanying note 8, *supra*.

gains from recontracting will outweigh the dynamic efficiency losses from permitting the reallocation of risks through modifications exacted and acceded to in large part because of limitations in the remedial system available to parties on breach. It should be added that efficient constraints on contractual modifications in all cases must apply not only to modifications *per se* but also to transactional substitutes (for example, rescission of the original contract and formation of a new contract, compromises of suits, and so on), so that these constraints are not reduced to largely formal significance.

In central respects, the problems entailed in formulating optimal modification rules are a function of legal limitations on remedies available for breach that prevent a non-breaching party from obtaining relief that puts him in the same position as if the contract had been performed. For the purposes of this paper, we have accepted these limitations as given, but without a re-evaluation of these limitations at least in some contexts (for example, the availability of specific performance,⁸⁶ the recovery of consequential damages,⁸⁷ the enforceability of penalty clauses⁸⁸), we are often left in a realm of second-best solutions to the question of the enforceability of contractual modifications.

As our analysis has attempted to show, neither a legal regime that enforces all contract modifications nor one that invalidates all modifications is efficient, apart from the savings in adjudication costs associated with a clear legal policy favouring one polarity over the other. The rather more complex and less definite regime suggested by our analysis, while generating higher adjudication costs than the polar choices, appears to weigh all the relevant economic variables. Because these involve difficult, or at least different, trade-offs from one class of case to the next, it is unrealistic to assume that a simple set of legal rules governing contract modifications is ever likely to evolve. The tangled history of the evolution of legal doctrine seems to reflect this reality. However, clear identification of the relevant variables seems a useful, albeit modest step towards reducing the conceptual confusion that has characterized this area of law.

⁸⁶ See Schwartz, *The Case For Specific Performance* (1979), 89 Yale L.J. 271.

⁸⁷ See Barton, *The Economic Basis of Damages For Breach of Contract* (1972), 1 J. Leg. Stud. 277; Birmingham, *Breach of Contract, Damage Measures, and Economic Efficiency* (1970), 24 Rutgers L. Rev. 273.

⁸⁸ See Goetz and Scott, *Liquidated Damages, Penalties and the Just Compensation Principle* (1977), 77 Colum. L. Rev. 554; Clarkson, Miller and Muris, *Liquidated Damages and Penalties: Sense or Nonsense?* (1978), 101 Wisc. L. Rev. 351; Rea, *The Efficiency Implications of Liquidated Damages and Penalty Clauses* (1984), 13 J. Leg. Stud. (forthcoming).