LEGAL REMEDIES AND REPUTATION AS SOLUTIONS TO MORAL HAZARD IN CONTRACTING

Thesis by

William Paul Rogerson

In Partial Fulfillment of the Requirements for

the Degree of Doctor of Philosophy

California Institute of Technology

Pasadena, California

1981

(Submitted on September 15, 1980)

ACKNOWLEDGEMENT

Financial support for the research and writing of this thesis was provided by an Earl C. Anthony Fellowship; a Shelby Cullom Davis Fellowship; a Canada Council Fellowship; and Graduate Research and Teaching Assistantships awarded by the California Institute of Technology.

The chairman of my committee, Louis Wilde, was a constant source of useful suggestions, guidance, and advice while I was writing my thesis. My principal advisors, Roger Noll and Louis Wilde, have had a large impact not only on my thesis, but also on my basic orientation towards the discipline in the four years I have been associated with them. I also wish to thank Robert Bates for the insightful comments and encouragement he provided through all stages of my work at Caltech. Alan Schwartz made a number of suggestions and comments which drastically improved my understanding of the phenomena I was investigating. I would like to thank Denise Noël for an excellent job of typing.

Most of all I would like to thank my parents for the love and support they provided me with while I pursued my Ph.D.

ABSTRACT

It is well known that noncooperative pursuit of individual gain may often result in a smaller total return than is otherwise possible. One purpose of contracts is to allow parties to commit themselves to courses of action which maximize their joint return. However, transaction and enforcement costs often create situations in which agreements between parties are either legally unenforceable or at least unenforceable in practice. This thesis explores solutions to this problem in two different contexts. In both cases it is assumed that the contract only specifies that an exchange will occur at a given price.

In the first case the value of the exchange to the buyer and/or seller is random at the time of contracting and the buyer has an opportunity to engage in expenditures prior to the date of the contract performance which will enhance the value of performance to him. These expenditures are called reliance. The legal institution which requires that a breaching party must pay the breachee an amount of money called damages, substitutes for an exhaustive contract specifying reliance behavior and behavior under all states. An unambiguous ranking of the six most common damage measures is obtained in terms of efficiency of the reliance decision.

In the second case, the quality of the product is variable. Buyers cannot judge the quality of the good they receive until after they consume it. Even then they may make mistakes in their judgments. This thesis derives the equilibrium quality distribution for goods produced and the equilibrium distribution of firms by the quality of good they produce, and identifies the specific factors which produce a reputation effect. Comparative statics allow analysis of the effects of restrictions on information flow and barriers to entry.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

As the actions of parties to some joint economic activity become more complex or difficult for third parties to monitor, two costs of contracting begin to grow larger. First, the enforcement and litigation costs begin to grow. Second, the transactions costs of drafting a contract which adequately describes the contemplated actions grow larger. The first cost produces the situation such that even if a contract is legally enforceable, it is in practice non enforceable because the size of litigation and enforcement costs relative to the gains from performance make the option of legal enforcement unattractive. The second cost tends to produce the situation where actions are not adequately specified to be legally enforceable. In both situations, parties to a contract find themselves in the position where they need not fear legal sanctions for breaching a contractual obligation.

The problem created by this is as follows. Suppose parties 1 and 2 contemplate a joint economic activity. Party 1 may well find that, given the latitude of choice allowed him, the action that maximizes his private return does not maximize the joint return of the two parties. In this case, both parties could be made better off by a side payment from party 2 to party 1 coupled with an enforceable contract specifying that party 1's actions be those which maximize the joint return. This is, however, by assumption impossible. The parties are thus prevented from realizing the greatest return from their joint activity. In the extreme case, inability to specify enforceable mutual obligations may prevent any joint activity at all.

Economists have pursued two lines of research related to this phenomenon. First, they have examined the nature of optimal "feasible" contracts. Given that certain aspects of the parties' behavior are not contractually specifiable, how close can the parties come to efficient behavior by contractual specification of those aspects of their joint activity for which this is possible? The principal-agent literature (see Shavell 1979 and Holmström 1979 for an overview) is an example of such an approach. In this case, the principal cannot directly monitor and therefore cannot contractually specify his agent's effort. However, basing the payment scheme from the principal to the agent on the output of the productive process in which the agent is engaged can allow the principal a certain amount of control over his agent's effort. The second perspective from which economists have considered the "nonenforceable contracts" question is in terms of non-contractual solutions to the problem. Can parties find institutional arrangements which allow them to organize their joint activity even in the absence of any contractual control over their mutual obligations? The most obvious example of a non-contractual solution is formation of a firm. (See Klein, Crawford, and Alchian 1978 for a discussion and literature survey of this subject.)

This thesis considers nonenforceable contracts from both perspectives. Chapter 3 adopts the former perspective. It considers a situation where the contract specifies that one party will sell a good or service to the other party at a future date and where the value of performance at the future date to the buyer might be much larger if he engages in some other activity ahead of time. For example, a rock promoter hiring a band can increase the value of the exchange to himself by advertising before the concert date. Expenses incurred prior to an exchange in anticipation of an exchange are called reliance. The buyer may be unwilling to engage in any reliance at all without some assurances that the other party will exchange at a previously agreed price. If negotiations over price occur after the buyer engages in reliance, he may be in a very weak negotiating position since he will lose money unless the seller performs. The buyer's expected return to reliance may be quite low or even negative in the absence of a contract which the buyer can negotiate prior to relying.

In the extreme case, this may mean that no exchange occurs, although both parties could have benefitted from it. More generally, the buyer may engage in less reliance than if he were assured of a particular price of exchange, and even though the exchange occurs, it does not generate nearly the aggregate value that it might have. Both parties could have been made better off had there been some manner of assuring one another's performance. A contract can do this.

Were it not for exogenous uncertainty, there would be no problem to analyze. The contract that an exchange will occur is clearly enforceable. Parties to the potential exchange would enter a contract if and only if a price exists which makes them both better off. The buyer maximizes his private return to reliance and the joint return to reliance simultaneously.

However, the case where exogenous uncertainty exists is not so simple. As examples of exogenous uncertainty, the cost of production may depend on the amount of rain that falls or future prices of inputs. In this case, the efficient solution typically involves some reliance but no exchange if the cost to the seller rises too high or the value to the buyer drops too low. A contract which maximizes joint value must now specify the complete set of contingencies and whether the exchange shall occur under each one. For reasons discussed earlier, such a contract is likely to be either impossible to enforce or extremely expensive to draft. It will be assumed that such a contract is not drafted. As well, it will be assumed that the extent of the buyer's reliance cannot be contractually specified for similar reasons. Contracts will be assumed to simply specify that an exchange will occur.

How, then, does a contract which simply specifies that an exchange will occur provide assurances of performance for the relier? The law could still (and does at times) provide for specific performance -- the relier could have the right to force the breacher to perform. More typically, a damage measure is embedded in the law which provides that a party to a contract who breaches must pay the breachee an amount of money called damages. These damage measures provide a measure of control over the seller's breach decision and buyer's reliance decision. Chapter 3 analyzes the incentives of the parties under the five most common damage measures and specific performance and obtains an unambiguous efficiency ranking of the six in terms of efficiency of the reliance decision. A number of policy prescriptions follow from this.

One of the most commonly cited factors which causes non-enforceable contracts to be honored is what economists have termed "reputation." Economic actors operating on the basis of self-interest may well honor non-binding contracts that appear not to be in their short run interest if future opportunities depend on adequate performance of current obligations. For example, the buyer from Chapter 3 may well choose a level of reliance which maximizes joint returns if doing so is a prerequisite for receiving future business. Chapter 4 analyzes the behavior of a market where reputation is the only incentive for fulfilling contracts. Sellers promise to deliver buyers a product of a specified quality. However, buyers cannot judge the quality of the good they receive until after they consume it. Chapter 4 derives the equilibrium quality distribution for goods produced and the equilibrium distributions of firms by the quality of good they produce, and identifies the specific factors which produce a reputation effect. Comparative statics allow analysis of the effects of restrictions on information flow and barriers to entry in the form of high fixed costs such as license fees and expensive training. These are often cited as characteristics of markets for professional services (Benham and Benham, 1975). A number of policy prescriptions follow from this.

REFERENCES

Benham, L. and Benham A. "Regulating the Professions: A Perspective on Information Control." Journal of Law and Economics 18 (1975):421-447.

Holmström, B. "Moral Hazard and Observability." Bell Journal of Economics 10 (Spring 1979):74-91.

Klein, Benjamin; Crawford, Robert G.; and Alchian, Armen, A. "Vertical Integration, Appropriable Rents, and the Competitive Contracting Process." Journal of Law and Economics 21 (1978):297-326.

Shavell, Steven. "Risk Sharing and Incentives in the Principal and Agent Relationship." Bell Journal of Economics 10 (Spring 1979):55-73.

CHAPTER 2

LITERATURE SURVEY

I CONTRACT REMEDIES

The following definitions are useful in reading the literature review. More complete definitions are contained in Chapter 3. Expenditures incurred by a party to a contract in anticipation of performance of the contract are called reliance. The institution of specific performance provides that parties cannot be unilaterally excused from performing contractual obligations. Damage measures allow parties to unilaterally breach contracts but provide that the breacher must pay the breachee an amount of money called damages. Expectation damages are the amount of money such that the breachee is made as well off as if the contract had been performed. Reliance damages are the amount of non-recoverable reliance expenses by the breachee. Liquidated damages are any specific sum of money agreed to by both parties in the contract as the penalty for breach.

Papers by Birmingham (1969), Barton (1972), and Shavell (1979) consider the problem of contract remedies from a game-theoretic standpoint. Birmingham (1969) was one of the first people to provide any economic analysis of contract remedies. He considers a prisoner's dilemma where each player has the option of honoring or breaching. While the honor-honor point maximizes joint returns and in fact maximizes each player's return, the breach-breach point is achieved since breach is a dominant strategy for both players. Birmingham points out that if a breaching player was forced to pay expectation damages to an honoring player, the honor-honor point is achieved because honor becomes the dominant strategy for each player. However, reliance damages, although resulting in honor-honor for some prisoners' dilemmas, result in honor-breach or breach-honor being achieved for some others. Birmingham therefore concludes that expectation damages result in efficient breach behavior while reliance damages result in too much breach.

Barton (1972) generalizes Birmingham's results to a broader class of two person games. His major contribution is to point out that a zero-sum game corresponds to the case where a market for substitute performance exists while a non zero-sum game corresponds to the case where no market for substitute performance exists. In a zero-sum game, the decisions of the players are irrelevant to efficiency considerations. Birmingham's conclusions are therefore relevant for the case of no market for substitute performance.

Shavell (1979) made the last major contribution to this literature by pointing out that the buyer's reliance decision as well as the seller's breach decision will be affected by the damage measure used. He demonstrates that reliance damages and expectation damages both produce reliance decisions larger than the efficient level and that reliance damages produce a less efficient reliance choice than expectation damages. Therefore both reliance and breach considerations suggest that expectation damages are more efficient than reliance damages.

At least three problems still exist with this literature. First, the assumption that inefficient breach behavior actually occurs may not be realistic. The literature has correctly identified that in the absence of further negotiations, some contract remedies such as reliance damages or specific performance create incentives for breach behavior which does not maximize joint returns. However, if post-contract negotiations occurred, a side payment coupled with efficient breach behavior could make both parties better off. The assumption that no such post-contract negotiations occur amounts to assuming that the transaction costs of such negotiations are larger than the efficiency gains. It is doubtful that this is always the case. Therefore it is important to consider the case where the buyer and seller do conduct post-contract negotiations to split the efficiency gains net of transaction costs. This turns out to affect the buyer's expected return to reliance and thus his reliance decision. Second, this literature has not considered specific performance, or such damage measures as restitution damages and liquidation damages. Furthermore, it has not formally modelled the case where no market exists and the buyer's reliance decision is too small. Third, the distortions on reliance in a market versus non-market setting have not been considered.

Another literature on the economic analysis of contracts has advocated more judicial acceptance of liquidated damages (Goetz and Scott, 1977) and specific performance (Kronman 1978; Schwartz 1979). Advocates for both measures use essentially the same arguments. Expectation damages do not protect the value of performance to the buyer if the buyer values a particular good much more than the average person values such goods or services. Goetz and Scott use the example of a diehard football fan who charters a bus to transport him and twenty of his friends to an important game. The bus company breaches and the fan and his friends are prevented from seeing the game. Although the loss in value to the fan may actually be huge, the courts have no way of objectively measuring this value and can only award the buyer the market evaluation of such services. Specific performance and liquidated damages can protect the buyer's full value and thus result in more efficient breach decisions.[1]

Courts also find it difficult to measure a buyer's search costs in locating a new source of supply, and therefore these are also typically excluded from expectation damages awards. Specific performance and liquidated damages can also protect this value and thereby produce more efficient breach decisions. Even when courts can calculate expectation damages, there is subjective uncertainty at the time of contracting concerning the court's estimate of the size of damages resulting from any potential breach. Specific performance and liquidated damages can remove this uncertainty as well as reduce costs of litigating over the size of expectation damages.

A further reason why parties might wish to use specific performance over expectation damages concerns the fact that specific performance transfers the property right for performance under large production cost rises from the seller to the buyer. The buyer may lack information about the probability of the seller wanting the breach while the seller might have very good information about this possibility and in fact consider it quite unlikely. In this case, the property right is worth much more to the buyer than the seller and specific performance allows transfer of the property right to the party which values it most. In some sense, the seller can communicate his high probability of performance to the buyer by offering specific performance at a small premium.

Kronman incorrectly argues on the basis of this point that specific performance should be used whenever no market for substitute performance exists. Schwartz correctly points out that availability of specific performance is irrelevant when a market for substitute performance exists and the above fact merely suggests a reason why parties with certain types of preferences and information might find specific performance to be mutually advantageous in the absence of a market for substitute performance. Chapter 3 contributes to this literature by producing an entirely new reason why parties to a contract might find specific performance or liquidation damages to be mutually advantageous. Both produce extremely efficient reliance decisions under the assumptions in the formal model of Chapter 3.

Diamond and Maskin consider contract remedies from a search theoretic perspective more common in the labor market literature (Diamond and Maskin, 1979a and 1979b). They consider a reason for contracting different from that of assuring performance. Buyers need a good on a specific date and must engage in costly search for a seller. A seller passed over cannot be returned to. Therefore, a buyer may agree to exchange with a particular seller but keep searching in hope of locating a better one. If this occurs, the buyer breaches his contract with the original seller and pays damages. Diamond and Maskin examine how various damage measures affect the amount of search. Potential exists to integrate this approach with the optimal reliance considerations of Chapter 3. Reliance is often conducted over time as opposed to occurring instantaneously. This distinction becomes important in a model such as Diamond and Maskin's, which explicitly considers the passage of time prior to exchange. The effect of various damage measures on the path of reliance relative to the efficient path of reliance thus becomes important. In particular, one might expect to find incentives for reliance to be speeded up relative to the efficient path in order to "force" the other party to honor.

The problem in Chapter 3, as with many topics in economics, can be viewed as a principal-agent problem. Shavell (1979) and Holmström (1979) provide state of the art treatments of the standard principal-agent problem. Their basic conclusion is that there is a tradeoff between risk allocation and incentive maintenance. In particular, if the agent is risk neutral, all problems vanish. An optimal contract can be written without exchange of any private information or joint calculations. The principal receives a fixed return and the remainder is the agent's. In the model of Chapter 3 even though both parties are risk neutral, an efficient contract is much harder to come by. It involves either exchange of private information and joint calculation during pre-contract negotiations (liquidated damages) or post-contract negotiations occurring costlessly and for the buyer to expect to capture all rents negotiated over (specific performance). The reason for this increased difficulty of achieving an efficient contract is that both parties are now one another's agent. The buyer makes a reliance decision and the seller makes a breach decision. Providing proper incentives for two agents simultaneously proves to be a more difficult problem.

II REPUTATION

At least two separate literatures deal with the mitigating effects of reputation on non-enforceability of contracts. One literature examines the micro-foundations of reputation from a repeated games standpoint. Radner (1979) considers the well known result that repeated games over a finite horizon tend not to exhibit "cooperative"[2] behavior because agents realize there is no gain to cooperation in the last period and this causes an unravelling of cooperative behavior backwards to the first period. Mathematically it seems an infinite horizon is needed for repetition of a game to produce "cooperative" behavior. Radner points out that agents content to come "very close to maximizing their return" may well demonstrate cooperative behavior in a finite horizon repeated game.

Rubinstein and Townsend (Rubinstein, 1979; Townsend, 1979) both consider repeated agency problems. In Rubinstein's problem, society must decide when to punish offenders, given that offenses may be committed accidentally in some cases. Society wishes to punish only those offenders who commit the offense purposefully. In Townsend's problem, an insurance company must decide on a claims policy when it only knows the probability distribution over whether the claimant will have a legitimate claim, but not whether any particular claim is legitimate. Both authors explicitly identify contracts which are efficient. In Rubinstein's model people never commit offenses purposefully; in Townsend's model all claims are legitimate. However, both results hinge very crucially on the assumption that the agent does not discount future income. The nature of the contracts in both cases is such that their optimality is extremely sensitive to the zero-discounting assumption. Furthermore, both models have aspects which make them inapplicable to the case of product quality. In Rubinstein's model, the principal can (and does, under the efficient contract) punish the agent periodically without severing the relationship. However, in the product quality case, the punishment itself consists of severing the relationship. Therefore, Rubinstein's model is inapplicable. In Townsend's model, the agent can (and does, under the efficient contract) report no claim, when in fact he has a legitimate claim. The efficient contract essentially allows the agent only the expected number of claims and it turns out to be to the agent's advantage to lie when necessary to keep his number of claims no greater than the allowed level. In the case of product quality, the firm often does not have total control over the consumer's judgment of product quality. Due to difficulties in judging product quality, the consumer's judgment may be somewhat random. Furthermore, the actual quality of the good produced may be subject to variance. Professional services, for example, exhibit both these characteristics. Therefore, the repeated games literature has not yet addressed formal models which capture the characteristics of the product quality problem.

The other area of research in the reputation field relevant to this thesis is more market-oriented research specifically addressing the product quality question. Akerlof (1970) points out that if firms cannot be differentiated by consumers and if consumers do not believe that a firm would continue to produce at its past quality level, the firms will all produce the lowest quality of goods. A market of "lemons" will result. Subsequent authors (Klein and Leffler, 1979; Dybvig and Spatt, 1980) adopt these necessary conditions for reputation and as well assume that consumers can judge quality with complete accuracy once they have received the good. Under this assumption, firms which misrepresent quality have a lifetime of one period. These papers are thus in a sense on the economics of "fly by night operators." However, a very large class of goods does not satisfy this assumption of perfect consumer accuracy. In particular, consumers are often capable of performing only very partial and vague evaluations of the quality of professional services they receive from doctors, lawyers, banks, mechanics, opticians, etc. Furthermore, the quality of a service from a given professional may vary from time to time. This combination of observer error and actual quality variance makes it difficult for consumers to correctly evaluate the quality of service that a firm produces. In such a market a producer can contemplate staying in the market over the long run even if he misrepresents the quality of his product. The quality elasticity of demand is no longer infinite, but a finite positive number. Chapter 4 analyzes such a market.

A paper by Schmalensee (1978) investigates the relationship between advertising and product quality. He characterizes the cases where product quality is positively or negatively correlated with the amount of advertising. In Schmalensee's analysis firms select their level of advertising but product quality is fixed. In the model of Chapter 4 of this thesis, firms select a level of quality. A natural extension of the model of Chapter 4 would be to allow firms the option of advertising.

Satterthwaite (1979) considers the market for "any product or service for which sellers' products are differentiated and consumers' search among sellers consists of inquiries to relatives, friends, and associates for recommendations" (Satterthwaite 1979, p. 483). He argues that an increase in the number of sellers in a market might paradoxically lead to a higher price. An increase in the number of sellers in a market will decrease the average amount of information a consumer has about any given doctor. If consumers only provide recommendations about doctors for which they possess some minimum amount of information, then a searching consumer will on average receive fewer recommendations from a consumer when the number of sellers increases. This lowering of the efficiency of search may decrease the price elasticity of demand which in turn raises market price. A large number of papers have considered the topic of price dispersion. See Wilde and Schwartz (1979) or Salop and Stiglitz (1977) for representative work and further literature. No progress, however, has been made in integrating the product quality literature and the price dispersion literature. This is clearly another area worth pursuing.

FOOTNOTES

1. The implicit assumption of these authors is that post-breach negotiation costs are very high.

2. The word "cooperative" may be misleading. The equilibrium concept used is still the Nash equilibrium, which is non-cooperative. Rather, the question is whether Nash strategies in the repeated games produce strategies which increase all players' period by period returns compared to the case where the game is played once.

REFERENCES

Akerlof, George A. "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism." Quarterly Journal of Economics 84 (August 1970):488-500

Barton, John H. "The Economic Basis of Damages for Breach of Contract." Journal of Legal Studies 1 (1972):277-304

Benham, L. and Benham, A. "Regulating the Professions: A Perspective on Information Control." Journal of Law and Economics 18 (1975):421-447.

Birmingham, Robert L. "Legal and Moral Duty in Game Theory: Common Law Contract and Chinese Analogies." Buffalo Law Review 99 (1969):103-110.

Darby, Michael R. "Qualitative Information, Reputation, and Monopolistic Competition." Mimeo, March 4, 1976.

__________, and Karni, Ed. "Free Competition and the Optimal Amount of Fraud." Journal of Law and Economics 16 (April 1973): 67-88.

Debreu, Gerard. Theory of Value, Cowles Foundation Monograph Number 17, New Haven and London: Yale University Press, 1959.

Diamond, P. and Maskin, E. "An Equilibrium Analysis of Search and Breach of Contract, I: Steady States." Bell Journal 10 (1979a):282-316

__________, and __________. "An Equilibrium Analysis of Search and Breach of Contract, II: A Non-Steady State Example." MIT Department of Economics Working Paper #237: 1979b.

Dybvig, Philip H. and Spatt, Chester S. "Does It Pay to Maintain a Reputation?" Mimeo, 1980.

Farnsworth, E. Allan "Legal Remedies for Breach of Contract." Columbia Law Review 70 (1970):1145-1216.

Goetz, Charles J. and Scott, Robert E. "Liquidated Damages, Penalties and the Just Compensation Principle: Some Notes on an Enforcement Model and a Theory of Efficient Breach." Columbia Law Review 77 (1977):554-594.

Green, Edward. "Agency, Information and Cooperation over Time." Mimeo, February 1980.

Holmström, B. "Moral Hazard and Observability." Bell Journal of Economics 10 (spring 1977): 74-91.

Joskow, Paul L. "Commercial Impassibility, the Uranium Market and the Westinghouse Case." Journal of Legal Studies 6 (1976): 119-176.

Klein, Benjamin; Crawford, Robert G.; and Alchion, Armen A. "Vertical Integration, Appropriable Rents, and the Competitive Contracting Process." Journal of Law and Economics 21 (1978):297-326.

Klein, Benjamin and Leffler, Keith B. "The Role of Price in Guaranteeing Quality." UCLA Discussion Paper No. 149.

Kronman, Anthony T. "Specific Performance." University of Chicago Law Review 45 (1978):351-382.

Kronman, Anthony T. and Posner, Richard A., eds., The Economics of Contract Law. Boston: Little, Brown and Company, 1979.

Landa, Janet and Grofman, Bernard "Games of Breach and the Role of Contract Law in Protecting the Expectation Interest." University of California, Irvine, School of Social Sciences, Research Report #34, June 1979.

Leff, Arthur Allen "Injury, Ignorance and Spite - The Dynamics of Coercive Collection." Yale Law Journal 80 (1970):1-29.

Macauley, Stewart. "Non-Contractual Relations in Business: A Preliminary Study." _American Sociological Review_ 28 (February 1963):55-69.

Nelson, Phillip. "Information and Consumer Behavior." _Journal of Political Economy_ 78 (1970):311-329.

Perloff, Jeffrey M. "Breach of Contract and the Foreseeability Doctrine of Hadley v. Baxerdale." Center for the Study of Organizational Innovation. University of Pennsylvania Discussion Paper #43, March 1979.

__________. "The Forward Market Effects of the Excuse Doctrines: Once More unto the Breach." Center for the Study of Organizational Innovation. University of Pennsylvania Discussion Paper #59, November 1979.

Radner, Roy. "Collusive Behavior in Noncooperative Epsilon-Equilibria of Oligopolies with Long but Finite Lives." Bell Labs Discussion Paper #144, August 1979.

Rothschild, Michael. "Models of Market Organization with Imperfect Information: A Survey." _Journal of Political Economy_ 81 (December 1973):1238-1308.

Rubinstein, Ariel. "An Optimal Conviction Policy for Offenses that May Have Been Committed by Accident." Hebrew University Department of Economics Research Report #114, January 1979.

Salop, S. "Information and Monopolistic Competition." American Economic Review 66 (May 1976):240-245.

__________, and Stiglitz, J. "Bargains and Ripoffs: A Model of Monopolistically Competitive Price Dispersion." Review of Economic Studies 44 (1977):493-510.

Satterthwaite, Mark A. "Consumer Information, Equilibrium, Industry Price, and the Number of Sellers." Bell Journal of Economics 10 (Autumn 1979):483-502.

Schmalensee, R. "A Model of Advertising and Product Quality." Journal of Political Economy 86 (1978)485-504.

Schwartz, Alan. "The Case for Specific Performance." Yale Law Journal 89 (1979):271-306.

Shavell, Steven. "Breach of Contract, Damage Measures and Resource Allocation." Mimeo, 1978.

__________. Risk Sharing and Incentives in the Principal and Agent Relationship." Bell Journal of Economics 10 (spring 1979): 55-73.

Townsend, Robert M. "Contract Length and the Discipline of Enduring Relationships." Mimeo, July 1979.

Wilde, Louis L. and Schwartz, Alan. "Equilibrium Comparison Shopping." Review of Economic Studies 46 (July 1979):543-553.

CHAPTER 3

EFFICIENT RELIANCE AND CONTRACT REMEDIES

I INTRODUCTION

The essential element of a contract is time: parties promise at some earlier date to perform specified actions at some later date. There are three broad classes of reasons people might want to enter such an arrangement. First, at the earlier date events out of control of either of the parties may still be uncertain. A contract can be a futures contract, transferring this exogenous uncertainty to those more willing to bear it. Second, at the earlier date events which can be affected by one of the parties may be subjectively uncertain to the other party. A contract can remove this endogenous uncertainty. The third reason is not associated with allocation of exogenous or removal of endogenous risk: it also applies to risk neutral people. The value of performance at the later date to one of the people might be much larger if he engages in some other activity ahead of time. For example, a rock promoter hiring a band can increase the value of the exchange to himself by advertising before the concert date. Expenses incurred prior to an exchange in anticipation of the exchange are called reliance. A party to an exchange may be unwilling to engage in any reliance at all without some assurances that the other party will exchange at a previously agreed price. By this it is not meant that the relier needs the variance of the price of exchange reduced. This falls under the second reason for contracting. Rather, he requires that its expected value be raised. If negotiations occur after the relier engages in reliance, he may be in a very weak negotiating position since he will lose money unless the other party to the exchange performs. The relier's expected return to reliance may be quite low or even negative in the absence of a contract which the relier can negotiate prior to relying.

In the extreme case, this may mean that no exchange occurs, although both parties could have benefited from it. More generally, the relier may engage in less reliance than if he were assured of a particular price of exchange, and even though the exchange occurs, it does not generate nearly the aggregate value that it might have. Both parties could have been made better off had there been some manner of assuring one another's performance. A contract can do this. I will call this third function of contracts "assuring performance."

Were it not for exogenous uncertainty and transactions and information costs, there would be no problem to analyze. In this case, a law that all contracts must be honored would induce efficient behavior. For the case of no exogenous uncertainty, parties to the potential exchange enter a contract if and only if a price exists which makes them both better off. The parties then maximize their own return by choosing levels of reliance which also maximize the aggregate value.

However, the case where exogenous uncertainty exists is not so simple. As examples of exogenous uncertainty, the cost of production may depend on the amount of rain that falls, or the buyer may be purchasing the good for resale and is uncertain of the future price. In this case, the efficient solution typically involves some reliance but no exchange if the cost of the seller rises too high or the value to the buyer drops too low. A simple legal provision that all contracts must be enforced does not induce this efficient behavior if the contract simply specifies that an exchange shall occur. Instead, the contract has to specify the complete set of contingencies and whether the exchange shall occur or not under each one.

In a world of zero transaction costs and costless gathering and processing of information, this is the "ideal" solution; parties to the contract guarantee that the exchange produces the maximum aggregate value and the negotiated price divides it between the two. However, drafting and, particularly, negotiating exhaustive contracts is expensive. The list of possible contingencies could be almost endless. Furthermore, a number of the contingencies may be private to one of the parties and very difficult to verify. This would allow the possibility of misrepresentation. For example, a seller's costs might rise enough that he would not want to exchange even though the contract specifies that he must; if his production process was complex he might easily be able to argue that costs had risen enough that the exchange should not take place according to the contract. The alternative would be to only specify contingencies external to the firm instead of using the cost variable. However, to do this both parties would essentially have to agree on what the firm's production function was; this is clearly an extremely costly process. In summary, to arrive at an efficient contract by this method would essentially amount to an exchange of all information and then joint calculation of an optimum. This process sacrifices the low cost, low information, and incentive-compatible properties of more decentralized decision making processes. Of course, some contingencies are important enough and easily verifiable enough that they are included in contracts; however, a large mass of contingencies are generally left unspecified in contracts.

How, then, does a contract which simply specifies that an exchange will occur provide assurances of performance for the relier? The law could still (and does at times) provide for specific performance -- the relier could have the right to force the breacher to perform. More typically, a damage measure is embedded in the law which provides that a party to a contract who breaches must pay the breachee an amount of money called damages. In either case, the institutions tend to provide assurances of performance by allowing the relier a private return to his reliance even in the event that breach is the efficient course of action. That is, the relier is insured to some extent against the possibility that his reliance may have no social return. As with many kinds of insurance, a moral hazard is created. In this case, the nature of the moral hazard is that the relier tends to over-rely. By solving the problem of under-reliance due to the lack of assurances of performance, we create a problem of over-reliance due to moral hazard.

The purpose of this paper is to compare the amount of moral hazard generated by different damage measures and by specific performance. The formal analysis is done in a particularly simple environment. I assume that the buyer and seller only specify in their contract that an exchange will occur at a fixed price. This is the situation encountered under high transaction and information costs. Only the buyer makes a reliance decision and only the seller's cost of production and the size of third party offers to purchase are subject to uncertainty at the time of contracting. The participants are assumed to be risk neutral in order to avoid confounding the analysis of efficient reliance and breach with that of efficient allocation of exogenous risk or removal of endogenous risk. I assume that both participants measure the value of the good to themselves in dollars. Together with the preceding assumptions, this means that participants will measure the value of a contract to themselves in expected value of dollars. If both parties are firms this assumption is fairly natural. See Rogerson (1980) for a discussion of the case where one of the parties is a consumer. This means that we can simply add the value of the contract to both players to obtain an efficiency index. A contract with a higher aggregate value is more efficient. Notice that the buyer and seller would always choose a more efficient institution over a less efficient one because in the former case they could negotiate a price which would make them both better off.

Shavell (1978) was the first person to point out that damage measures might distort the reliance decision. The analysis of this paper owes a debt to that of Shavell, but is substantially different for the following reasons. First, the agents are allowed to negotiate away potentially inefficient breach behavior at the given level of reliance. Modelling this process requires that the relier form expectations over payoffs resulting from negotiations that will occur at a future time. (This same model is also used to formally demonstrate that reliance is in general too small without assurances of performance.) Second, third party offers to purchase are considered separately from costs of production. This allows a distinction to be drawn between restitution and expectation damages. As well, it allows a clearer picture of the information requirements for the various institutions. Third, a distinction is made between the case where a market for substitute performance exists and the case where no market for substitute performance exists. The ranking of institutions is substantially affected by this factor. Fourth, liquidated damages, restitution damages, and specific performance are considered, as well as expectation damages and reliance damages.

II THE MODEL

The buyer of the good intends to use the good as an input in some production process and to sell the result or to consume the good himself. In either case he can engage in reliance, r. The value of the good to him is then v(r). Therefore, the net value of the good to him is

$$v(r) - r. \tag{1}$$

If the buyer engages in reliance and no exchange takes place, he may be able to obtain some scrap value for the reliance or, if he is not so lucky, may have to pay a disposal cost. Let $\bar{v}(r)$ denote this amount. Therefore the value of no exchange to him is:

$$\bar{v}(r) - r. \tag{2}$$

Assume that:

$$\begin{aligned}
&\text{(i)} && v \text{ and } \bar{v} \text{ are defined and continuous over } [0,\infty),\\
&\text{(ii)} && v(r) - r \text{ has a unique global maximum at } r_e,\\
&\text{(iii)} && v(r) - \bar{v}(r) \text{ is nondecreasing},\\
&\text{(iv)} && \bar{v}(r) - r \text{ is decreasing.}
\end{aligned} \tag{3}$$

Assumption (ii) means that even if one is certain the good will be produced and exchanged, eventually some optimal level of reliance is reached; past some point the net return to reliance begins to decline. The assumption of uniqueness is made for technical convenience. Assumption (iii) means that engaging in reliance for its primary purpose is at least as profitable as engaging in reliance for its scrap value. Assumption (iv) means that engaging in reliance solely to sell it as scrap is unprofitable; the buyer would only engage in reliance if there were some hope of purchasing the good. Note in particular that no convexity assumptions need to be made.

The seller produces the good at a cost of c. This cost varies randomly; the realization of the random variable is unknown at the time r is chosen. Let k be the best other offer that the seller of the good receives for his good between the time the contract is entered and the time that the exchange is supposed to occur. If a market for the good exists, k is the market price at the time of exchange. The realization of k is unknown at the time the contract is entered and the reliance decision is made. Assume without loss of generality that c and k are defined over the probability space [0,1] with Lebesgue measure. Let θ denote an element of [0,1] . Assume that c is always non-negative and k is bounded from below. This simply means that it costs at least zero to produce the good and the market price is bounded from below.

If the good in question is not unique and the buyer of the good intends to use the good as an input in some production process and to sell the result one might expect that a rise in the price, k, would also cause a rise in the price in the market for the buyer's output. As a consequence v(r) would also rise. In this case, interpret k as the net increase in price in the input market after the price increase in the output market is accounted for. Therefore, the same mathematics applies to this more complex case. For ease of exposition I will address the simpler case where v is not random.

The buyer and seller negotiate a contract which states that exchange will occur at a price, p. Six types of contract enforcement institutions will be considered. Specific performance is the simplest. The buyer has the right to demand that the exchange occur at p. Under reliance damages, the breacher must compensate the relier for all non-recoverable reliance expenditures. Therefore the seller must pay the buyer $r - \bar{v}(r)$ if the seller breaches.

The definition of expectation damages is more complex. This is the damage measure currently in general use in the courts. Its intention is to put the relier in the same financial position as if the contract were carried out. Suppose first that a competitive market exists where substitute performance can be purchased. If the buyer uses the good himself he would receive

$$v(r) - r - p. \tag{4}$$

However, if he sold to the highest alternative buyer he would receive

$$\bar{v}(r) - r - p + k. \tag{5}$$

Therefore, assuring the buyer of a level of profits equal to that he would have received had performance occurred amounts to assuring him the maximum of (4) and (5). If the breaching seller pays the buyer

$$k - p \qquad (6)$$

the buyer can decide which of (4) or (5) to receive by deciding whether or not to purchase substitute performance.

In many cases, however, there is no market for substitute performance. The most obvious case is where the good in question is unique. However, at least two other cases also arise. Sometimes there are many types of the good but the reliance is only useful for one type. For example, a rock promoter could hire any band but once he has hired Abba and advertised that Abba is coming and sold tickets to an Abba concert, some other rock band cannot provide substitute performance. Time is sometimes a crucial variable. If the seller breaches too close to the date of performance and it is crucial that the performance occur precisely at that time, no substitute performance may be obtainable on such short notice. In the extreme, the breacher may not indicate that he is breaching until the moment of expected performance. Notice that when time plays this role specific performance is not a possibility.

Expectation damages when no market for substitute performance exists are defined by

$$v(r) - \overline{v}(r) - p \qquad (7)$$

When these are paid to the buyer (who sells his reliance for $\overline{v}(r)$) the buyer's net return is

$$v(r) - r - p \qquad (8)$$

which is what he would have received had he received the good and used it himself.

This definition of expectation damages for the case of no market for substitute performance is the one the courts generally use. However, when $k > v - \bar{v}$, this rule may result in the situation where the seller breaches and sells the good to the third party for k. The seller's net profit over the case where he sells to the original buyer is $k - v(r) - \bar{v}(r)$. If one believed that the original buyer could also have resold the item for k to the third party, then protecting his expectation interest requires that damages be the maximum of (7) and $k - p$, where k is now the price at which the breaching seller sold to a third party. I will call this variant of expectation damages ideal restitution damages.

I call this variant "ideal" restitution damages because although it will be seen to possess good properties, the courts could probably not administer it. To do so, the courts would have to be able to determine the value of the highest third party offer to purchase. Such a process would generally be prohibitively expensive and error-prone, if not impossible. I will call a more feasible variant of this damage measure restitution damages. Under restitution damages the seller pays expectation damages if he elects to sell to no one. However, if he elects to sell to a third party he pays the maximum of $k - p$ and expectation damages. Note that under both ideal restitution damages and restitution damages the seller never has an incentive to sell to a third party.

Liquidated damages is the last damage measure that will be considered. Under this measure, parties specify in the contract a sum of money which the seller must pay the buyer in the event of breach.

III A MARKET FOR SUBSTITUTE PERFORMANCE EXISTS

We measure the efficiency of the outcome by summing the value to the buyer and seller. Three decisions affect the aggregate value generated--the amount of reliance engaged in by the buyer; whether or not the seller produces the good; and whether the buyer consumes the good produced by the seller, some other good or no good at all. Let P_1 be those values of θ such that the seller produces the good and the buyer consumes it or some other good. Let P_2 be those values of θ such that the seller produces the good and the original buyer does not consume a good. Let P_3 be those values of θ such that the seller does not produce the good and the buyer buys a good from someone else. Let P_4 be all other values of θ. Then the aggregate expected value of the exchange can be written as a function of r, P_1, P_2, P_3, and P_4. Let λ denote Lebesgue measure. Let F be the function determining aggregate value of the exchange.

$$F(r,P_1,P_2,P_3,P_4) = \int_{P_1} v(r) - r - c(\theta)d\lambda$$

$$+ \int_{P_2} \overline{v}(r) - r + k(\theta) - c(\theta)d\lambda$$

$$+ \int_{P_3} \bar{v}(r) - r - k(\theta) d\lambda$$

$$+ \int_{P_4} \bar{v}(r) - r d\lambda$$

Proposition 1:

Define the following four sets:

$$P_1^*(r) = \{\theta : c(\theta) \leq k(\theta) \wedge v(r) - \bar{v}(r) \geq k(\theta)\} \tag{11}$$

$$P_2^*(r) = \{\theta : c(\theta) \leq k(\theta) \wedge v(r) - \bar{v}(r) < k(\theta)\} \tag{12}$$

$$P_3^*(r) = \{\theta : c(\theta) > k(\theta) \wedge v(r) - \bar{v}(r) \geq k(\theta)\} \tag{13}$$

$$P_4^*(r) = \{\theta : c(\theta) > k(\theta) \wedge v(r) - \bar{v}(r) < k(\theta)\} \tag{14}$$

For any $r \in [0,\infty)$, $P_1^*(r), P_2^*(r), P_3^*(r)$, and $P_4^*(r)$ uniquely (up to inclusion or exclusion of sets of measure zero) maximize $F(r, P_1, P_2, P_3, P_4)$.

Proof:

Obvious. □

The intuition of Proposition 1 is clear. It is efficient to produce if and only if the cost of production is less than or equal to the market price. It is efficient to consume if and only if the marginal benefit from consumption is greater than or equal to the market price. Define f(r) as the function yielding an aggregate value for r when P_1, P_2, P_3, and P_4 are chosen optimally:

$$f(r) = F(r, P_1^*(r), P_2^*(r), P_3^*(r), P_4^*(r)) \qquad (15)$$

<u>Proposition 2</u>:

The function f(r) is continuous on $[0,\infty)$. Furthermore $f(r_e) > f(r)$ for all $r > r_e$. Therefore f achieves a global maximum and all such global maxima occur in $[0,r_e]$.

<u>Proof</u>:

See Appendix. □

Therefore an optimum level of reliance exists. Let $\mathcal{R}$ be the set of all such levels of reliance. (The maximum may not be unique.) By referring to the proof it is easy to see that r_e will generally not be an element of $\mathcal{R}$. This is true, for example, if f is differentiable. Loosely speaking, the element r_e will be in $\mathcal{R}$ when v exhibits a large kink at r_e. The most interesting observation to be made about the case where a market for substitute performance exists is that in the absence of any contract institution, the buyer and seller will act optimally.

<u>Proposition 3</u>:

When a market for substitute performance exists, the buyer and seller act optimally in the absence of a contract institution.

Proof:

See Appendix. □

The proof is simple. The buyer will consume the good if and only if $v(r) - \bar{v}(r) \geq k$; otherwise he is better off reselling it. The seller will produce the good if and only if his costs of production are less than or equal to the market price. The only real question is whether the buyer relies optimally. He does so because he experiences all the marginal social costs and benefits when he varies r. Therefore risk neutral parties do not need a contract institution when a market exists. There is no need for the relier to be assured of performance for optimal reliance to occur because his negotiating position is not damaged when he relies. He can purchase at the market price regardless of his level of reliance.

Even though no contract institution is needed to assure performance in the case where a market for substitute performance exists, it is still interesting to determine how much moral hazard exists, because a contract institution may be in use for one of the other two reasons outlined in the introduction. Expectation damages are easiest. The buyer and seller consume and produce optimally (i.e., the buyer consumes if and only if $v(r) - \bar{v}(r) \geq k$ and the seller produces if and only if $c \leq k$). The buyer can view himself as always receiving the good at price p. (If the seller honors, this is automatically so. If the seller breaches, the buyer receives enough money so he need pay only p more dollars to purchase the good. If the buyer breaches, he must pay enough money so it would still cost him in total p dollars to buy the good or the market.) The buyer relies optimally in this case. Specific performance is equivalent to expectation damages when a market for substitute performance exists because the seller will purchase a good on the market to fulfill his contract obligation if $c > k$. Therefore specific performance also results in efficient behavior. Proposition 4 states these results.

Proposition 4:

When a market for substitute performance exists, the buyer and seller act optimally using either expectation damages or specific performance.

Proof:

See Appendix. □

Therefore, when a market for substitute performance exists, most of the problems of concern in this paper vanish. The relier needs no assurances of performance to rely efficiently. The institutions of expectation damages and specific performance produce no moral hazard and result in efficient allocation of resources. In the next section it will be seen that both problems exist in the absence of a market for substitute performance.

IV. NO MARKET FOR SUBSTITUTE PERFORMANCE EXISTS

A. Efficient Behavior

Once again we measure the efficiency of the outcome by summing the value to the buyer and seller. The same three decisions affect the aggregate value generated, except that now the buyer does not have the option of consuming some other substitute good if the seller breaches. Let P_1 be those values of θ such that the good is produced and the buyer consumes it. Let P_2 be those values of θ such that the good is produced and a third party consumes it. Let P_3 be those values of θ such that the good is not produced. Then the aggregate value of the exchange can be written as a function of r, P_1, P_2 and P_3. When possible, I will re-use the same symbols that were used for the corresponding expressions in section III since the analysis is very similar. Let F denote the aggregate expected value of an exchange.

$$F(r,P_1,P_2,P_3) = \int_{P_1} v(r) - r - c(\theta)d\lambda + \int_{P_2} \overline{v}(r) - r + k(\theta) - c(\theta)d\lambda + \int_{P_3} \overline{v}(r) - rd\lambda \qquad (16)$$

Define the following three sets:

$$P_1^*(r) = \{\theta \ : c(\theta) \le v(r) - \bar{v}(r) \ \wedge v(r) - \bar{v}(r) \ge k(\theta)\} \tag{17}$$

$$P_2^*(r) = \{\theta \ : c(\theta) \le k(\theta) \ \wedge k(\theta) \ge v(r) - \bar{v}(r)\} \tag{18}$$

$$P_3^*(r) = \{\theta \ : c(\theta) > v(r) - \bar{v}(r) \ \wedge c(\theta) > k(\theta)\} \tag{19}$$

Proposition 5:

For any $r \in [0,\infty)$, $P_1^*(r)$, $P_2^*(r)$, and $P_3^*(r)$ uniquely (up to inclusion or exclusion of sets of measure zero) maximize $F(r,P_1,P_2,P_3)$.

Proof:

Obvious. □

Figure 1 illustrates Proposition 5.

FIGURE I

AGGREGATE RETURN TO RELIANCE

θ is such that the largest of $\{v - \bar{v}, c, k\}$ is	Joint value
$v - \bar{v}$	$v - r - c$
k	$k - \bar{v} - r - c$
c	$\bar{v} - r$

If θ is such that the most profitable course of action for the buyer and seller is to honor the contract, then reliance is useful up to r_e.

However, if θ is such that the most profitable course of action is for a third party to consume the good or for no production to occur then, ex post, reliance always exhibits a negative return. Therefore, no matter what happens, it is never profitable to rely past r_e. Define f(r) as the function yielding aggregate value for r when P_1, P_2, and P_3 are chosen optimally:

$$f(r) = F(r, P_1^*(r), P_2^*(r), P_3^*(r)). \tag{20}$$

Let $\mathcal{R}$ be the set of values for r which maximize f.

Proposition 6:

The function f(r) is continuous on $[0,\infty)$. Furthermore, $f(r_e) > f(r)$ for all $r > r_e$. Therefore, f achieves a global maximum and all such global maxima occur in $[0,r_e]$.

Proof:

The proof is similar to that of Propsition 2. □

Therefore an optimum level of reliance exists. Let $\mathcal{R}$ be the set of all optimal levels of reliance. As for Proposition 2, r_e will generally not be an element of $\mathcal{R}$ unless v exhibits a large kink at r_e.

B. Negotiations

In the analysis that follows, it is necessary on four separate occasions to model the process by which the buyer forms expectations over a future uncertain payoff to be determined by negotiations between the buyer and seller. It seems most economical to present a model of this process at the outset. The four situations where the buyer needs to form these expectations in order to make his reliance decision are as follows. The first instance occurs if no contract is entered. The buyer may still choose to rely to some extent and then negotiate a price with the seller at the time of exchange. The other three instances occur when a contract institution would produce inefficient breach behavior conditional on the level of reliance. It seems likely that a negotiation involving a side payment would remove any such inefficiency. This situation occurs for specific performance, reliance damages, and restitution damages. (Ideal restitution damages, expectation damages, and liquidated damages result in efficient behavior conditional on the level of reliance without post contract negotiations.)

In all cases, there is a natural upper and lower bound on the size of the side payment determined by what the agents could secure for themselves in the absence of cooperative action. The agents can therefore be viewed as essentially negotiating over how to divide up a sum of money -- the difference between the upper and lower bound. I assume that there exists a number α in the interval [0,1] such that the buyer expects to receive α of the sum of money. The number α is thus an index of negotiating strength. As α grows larger the buyer feels that he will secure more of the rent from any negotiation.

It should be stressed that the rankings of damage measures hold even when different values of α hold for different damage measures. This is true for the following reason. Select any two of the damage measures. Call them measure 1 and measure 2. Let r be in R_i if and only if there exists an α in [0,1] such that r would be chosen by the buyer under damage measure i and that value of α. Then either $\sup R_1 \leq \inf R_2$ or $\sup R_2 \leq \inf R_1$. That is, the sets of possible reliance choices are disjoint and one is larger than the other. This allows a ranking of the efficiency of the reliance choice even if the buyer uses a different value of α under different damage measures. Figure 6 in the summary to Section IV illustrates this point. The discussion surrounding Figure 6 presents some technical qualifications to this intuition.

Under the assumption that post-contract negotiations always occur to remove inefficiencies at the given reliance level the size of α reflects the relier's expectations of his strength in the upcoming negotiations. However, these negotiations may not always result in efficient behavior due to bluffing or poor information, for example. As well, these negotiations are costly and thus consume some of the returns they generate. In this more general context a lower value of α can be generated by a higher expected negotiation cost or higher probability of arriving at no agreement. I will formally refer to α merely as an indicator of negotiating strength for ease of exposition.

This model of negotiations performs two functions in the later analysis. First, it establishes that the expected payoff from negotiations to the relier depends on his financial position in the absence of cooperation. This latter variable is of course affected by his reliance decision. Second, it establishes that the first effect is manifested in a smooth, regular fashion. In particular, if the buyer can increase his expected return in the absence of cooperation, he will also increase his return from negotiations.

When the seller has an incentive to make an inefficient breach or sales decision at the given level of reliance, the buyer's return to reliance is the same if one assumes that no post-contract negotiations occur or if one assumes that post-contract negotiations occur but that the buyer receives none of the increase in joint returns (i.e., α equals 0). However, these two cases are not equally efficient. At least when there are no transactions costs and where post-contract negotiations always produce agreement, the case of post-contract negotiations and α equals zero produces a joint return of f(r), while the case of no post-contract negotiations produces a smaller joint return. Therefore, to compare efficiency of institutions by comparing the values f assumes under the reliance decisions they generate, requires the assumption that post-contract negotiations occur to the same extent and at the same cost under all the institutions.

C. Behavior with No Contract

If no contract is entered, the buyer must first choose a level of reliance. The realization of θ then occurs and the buyer and seller negotiate a price at which to exchange, if they exchange. The negotiated price must always be greater than or equal to the maximum of $c(\theta)$ and $k(\theta)$ and less than or equal to $v(r) - \overline{v}(r)$.

If the price was below $c(\theta)$, the seller would find it more profitable to not produce. If it was below $k(\theta)$ the seller would find it more profitable to sell to the third party. If it was above $v(r) - \overline{v}(r)$, the buyer would be better off simply be selling his reliance for scrap. Therefore, in terms of the framework in Section B, the buyer expects to exchange if and only if

$$\max\{c(\theta),k(\theta)\} \leq v(r) - \overline{v}(r) \tag{21}$$

in which case he expects to receive

$$\alpha(v(r) - r - \max\{c(\theta),k(\theta)\}) + (1 - \alpha)(\overline{v}(r) - r). \tag{22}$$

Let $b(r,\alpha)$ be his expected return from reliance, r, given his subjective expectations, α.

$$b(r,\alpha) = \int_{P_1^*(r)} \Big\{\alpha(v(r) - r - \max\{c(\theta),k(\theta)\}) + (1 - \alpha)(\overline{v}(r) - r)\Big\} d\lambda + \int_{P_2^*(r) \cup P_3^*(r)} \overline{v}(r) - r d\lambda. \tag{23}$$

Proposition 7 shows that the reliance decision is, in general, too small. Furthermore, the extent to which the reliance decision is inefficient depends in a monotonic fashion on how confident the buyer is of his negotiating strength. If the buyer has absolutely no confidence in his negotiating strength, he engages in no reliance. This may mean that no exchange ever takes place. As the buyer's confidence increases, his reliance decision grows until, finally, if he expects to receive the maximum possible in all negotiation situations, he relies efficiently. Furthermore, not only his level of reliance but also its efficiency grow monotonically with his confidence level. (This is true even though f may not be monotonically increasing in r over $[0,r_e]$.)

Comparison of Figure 1, which diagrams the aggregate returns to reliance, and Figure 2, which diagrams the buyer's returns to reliance under no contract provides intuition into these results.

FIGURE 2

BUYER'S RETURN TO RELIANCE UNDER NO CONTRACT

θ is such that the largest of $\{v - \bar{v}, c, k\}$ is	Value to Buyer
$v - \bar{v}$	$\alpha[v - r - \max\{c,k\}] + (1 - \alpha)[\bar{v} - r]$
k	$\bar{v} - r$
c	$\bar{v} - r$

The marginal return to reliance when k or c is largest is the same under both schemes. However in Figure 2 the marginal return to reliance is less when it is optimal for the buyer to receive the good. It is less to the extent that the seller can bargain away the rents associated with performance. Since the return to reliance is smaller, the buyer relies less.

Proposition 7:

The function $b(\cdot,\alpha)$ achieves its supremum. Let $N(\alpha)$ be the set of all values for r such that $b(\cdot,\alpha)$ is maximized. Then

(i) $N(1) = R$

(ii) $N(0) = \{0\}$

(iii) If $\alpha_1 < \alpha_2$ then $\sup N(\alpha_1) \leq \inf N(\alpha_2)$.

(iv) If $\alpha_1 < \alpha_2$ then $\sup \{f(r):r \in N(\alpha_1)\} \leq \inf \{f(r):r \in N(\alpha_2)\}$

(v) N is an upper hemi continuous correspondence.

Proof:

See Appendix. □

It is interesting to note that Proposition 7 does not depend on any concavity assumptions concerning v. In particular, even though f may have numerous local maxima and non-concavities, as α goes up, $f(N(\alpha))$ also goes up (Proposition 7: iv). The correspondence N "passes over" values of r such that f(r) is decreasing. The entire proof is driven by the assumption that $\bar{v}(r) - r$ is decreasing. See Lemma 3 in the Appendix for an explanation of the nature of the proof. This observation can be made about the succeeding propositions as well, but will only be made here for economy of presentation.

D. Behavior Under Expectation Damages

The buyer always receives $v(r) - r$. If the seller breaches the buyer simply receives a net of $v(r) - r$ dollars. If the seller produces, he gives the good to the buyer if and only if $v(r) - \bar{v}(r)$

$\leq k(\theta)$. Otherwise he sells to a third party and pays the original buyer damages. Therefore, the buyer never has a resale opportunity and he receives $v(r) - r$ if the contract is honored. Recall that the unique global maximum to $v(r) - r$ occurs at r_e. The buyer clearly chooses r_e under expectation damages. By Proposition 6, the reliance choice of the buyer is generally larger than the efficient level of reliance. This observation was first made by Shavell (1978). The buyer is insured against both third party offers and cost rises which render performance inefficient. Even though there is no social return to reliance in these cases, the buyer receives a private return to this reliance. Not surprisingly, the buyer thus over-relies.

E. Behavior under Reliance Damages -- No Post-Contract Negotiations

The buyer receives

$$\max\{v(r) - r - p,\ k(\theta) + \bar{v}(r) - r - p\} \tag{24}$$

if the seller honors the contract, since the buyer has the option of reselling to a third party, and 0 if the seller breaches. The seller receives

$$p - c(\theta) \tag{25}$$

if he honors the contract,

$$k(\theta) - c(\theta) - r + \bar{v}(r) \tag{26}$$

if he sells to a third party, and

$$-r + \bar{v}(r) \tag{27}$$

if he does not produce.

Now consider the following three expressions:

$$p + r - \bar{v}(r) \tag{28}$$

$$c(\theta) \tag{29}$$

$$k(\theta) \tag{30}$$

From the above, the seller honors the contract if (28) is the largest of the three; he does not produce if (29) is the largest of the three; he sells to the third party if (30) is the largest of the three. If two or more of the terms are tied for largest, he is indifferent between the actions associated with them. I will assume that if (28) is involved in a tie for the largest, the seller honors. If (29) and (30) are tied for the largest, I will assume that the seller sells to the third party. Let $H(r,p)$ be all values of θ such that the seller honors the contract. Let $B_1(r,p)$ be the values such that the seller sells to the third party and let $B_2(r,p)$ be the values such that the seller does not produce. Let $B(r,p) = B_1(r,p) \cup B_2(r,p)$.

Now suppose that $v(r) > p + r$. Then it is clear that

$$H(r,p) \subseteq P_1^*(r)$$

$$B_1(r,p) \supseteq P_2^*(r)$$

$$B_2(r,p) \supseteq P_3^*(r). \tag{31}$$

If $v(r) = p + r$, then

$$H(r,p) = P_1^*(r)$$

$$B_1(r,p) = P_2^*(r)$$

$$B_2(r,p) = P_3^*(r). \tag{32}$$

If $v(r) < p + r$, then

$$H(r,p) \supseteq P_1^*(r)$$

$$B_1(r,p) \subseteq P_2^*(r)$$

$$B_2(r,p) \subseteq P_1^*(r) \tag{33}$$

That is, if the buyer makes a profit from relying and receiving the good, the seller does not give the good to him often enough. If the buyer makes zero profits from relying and receiving the good, the seller acts efficiently. If the buyer makes negative profits from relying and receiving the good, then the seller honors the contract more often than is efficient.

In the following paragraph I show that the buyer will never choose an r and p such that $v(r) - r - p < 0$, because if this is true then the buyer makes at best expected profits of zero and generally makes negative expected profits. Suppose that $v(r) - r - p < 0$. Then if $\theta \in H(r,p)$, we also know that $k(\theta) + \bar{v}(r) - r - p \leq 0$ so that the buyer at best makes zero profits. If $\theta \in B_1(r,p) \cup B_2(r,p)$ the buyer makes zero profits. Therefore the buyer makes at best zero profits and there is no incentive for him to have entered the contract.

Because of the observation in the last paragraph, it is reasonable to assume that at the value of p chosen by the buyer and seller the set of values for r such that $v(r) - r - p \geq 0$ is nonempty. By the above, the buyer also chooses his reliance level from this set. The buyer's expected return to reliance is

$$\int_{H(r,p)} \max\{v(r) - r - p, k(\theta) + \bar{v}(r) - r - p\} d\lambda + \int_{B(r,p)} 0 d\lambda \tag{34}$$

However, if $\theta \in H(r,p)$, then $k(\theta) + \bar{v}(r) - r - p \leq 0$, because (30) is less than or equal to (28). Therefore, we can rewrite (34) as

$$a(r,p) = \int_{H(r,p)} v(r) - r - p d\lambda \tag{35}$$

over the domain where (34) is non-negative. As well, (34) is negative if and only if (35) is negative. Therefore r^* maximizes (34) if and only if r^* maximizes (35).

The deduction that the buyer chooses r to maximize $a(r,p)$ allows the fairly immediate conclusion that the buyer's choice of reliance will be at least as large in this case as for the case of expectation damages. The buyer receives $v(r) - r - p$ when the seller honors, just as in the expectation case. However, now the buyer receives nothing if the seller breaches. Therefore he has an incentive to choose a larger r to encourage the seller to honor more often. Analogous to the previous propositions, because of the generality of the assumptions the formal statement of the proposition allows the possibility that the reliance choice under reliance damages will equal that under expectation damages. However, this will only happen in special cases such as where $v(r) - r$ has a large "kink." Generally, the reliance choice under reliance damages will be larger.

Proposition 8 (Shavell):

Suppose that there exists an r such that $v(r) - r - p \geq 0$. Let $A(p)$ be the set of values of r which maximize $a(r,p)$. Then

(i) $r \in A(p) \Rightarrow r \geq r_e$ for every p.

(ii) If there exists an $\bar{r}$ such that $v(r) - r < 0$ for every $r \geq \bar{r}$, then $A(p) \neq \emptyset$ for every p.

Proof:

See Appendix. □

Using expressions (31)-(33) and the fact that $v(r) - r - p \geq 0$ at the chosen level of reliance, the seller's choices are biased away from efficiency at the given level of reliance in the following fashion. The seller will sometimes not produce when it would have been more efficient to produce and honor the contract. He will sometimes produce and sell to the third party when it would have been more efficient to honor the contract. However, he produces and sells to a third party if and only if this is more efficient than not producing. Therefore the efficiency of reliance level r is less than or equal to $f(r)$, which is the efficiency achieved if the seller and buyer act optimally given the reliance level r. To show that the institution of reliance damages is no more efficient than that of expectation damages, it is thus sufficient to show that

$$f(r_e) \geq \sup\{f(r): r \in A(r)\}. \tag{36}$$

However, this is automatically true by Proposition 6 since $r \in A(p)$ implies that $r \geq r_e$.

Proposition 9 (Shavell):

Under the assumption of no post-contract negotiations, the institution of reliance damages is no more efficient than that of expectation damages.

Proof:

As above. □

F. Behavior under Reliance Damages -- Post-Contract Negotiations

In the last section the buyer assumed that the seller's breach behavior would not be efficient at the given level of reliance. To maximize his own return the seller sometimes does not honor the contract when doing so would actually increase the joint return of the buyer and seller. In this situation the potential exists for the buyer and seller to negotiate a side payment from the buyer to the seller in return for the seller honoring the contract. I call these negotiations "post-contract negotiations." The most interesting effect of these negotiations for this analysis is that the relier takes them into account when making his reliance decision, and in general this changes his reliance decision.

It is once again possible to show that if $v(r) - r - p < 0$ the buyer makes at best zero expected profits and generally makes negative expected profits. I will not present the proof since it is essentially the same as the analogous proof in section E. As in section E, I assume that at the value of p chosen by the buyer and seller there exists an r such that $v(r) - r - p \geq 0$. The buyer always chooses an r such that $v(r) - r - p \geq 0$. The expected return to reliance over this domain is

$$a^*(r,p,\alpha) = \int_{H(r,p)} v(r) - r - p d\lambda$$
$$+ \alpha \int_{P_1^*(r) \cap B_1(r,p)} v(r) - \bar{v}(r) - k(\theta) d\lambda$$
$$+ \alpha \int_{P_2^*(r) \cap B_2(r,p)} v(r) - \bar{v}(r) - c(\theta) d\lambda \qquad (37)$$

The return to reliance under no post-contract negotiations, (35), differs from (37) in that (37) has the extra term

$$\alpha \int_{P_1^*(r) \cap B_1(r,p)} v(r) - \bar{v}(r) - k(\theta) d\lambda$$
$$+ \alpha \int_{P_2^*(r) \cap B_2(r,p)} v(r) - \bar{v}(r) - c(\theta) d\lambda.$$

The marginal return to increased reliance from this extra term may be positive or negative and as a consequence it cannot be stated in general whether reliance under post-contract negotiation is smaller or larger than reliance under no post-contract negotiations. However one observation can be made. Reliance is still at least as large as r_e. Shavell's original observation that reliance damages produce an overly large reliance decision under no post-contract negotiations thus generalizes to the case of post-contract negotiations. By Proposition 6, therefore, reliance damages under post-contract negotiations produce a less efficient outcome than expectation damages. The intuition for this result is similar to that for Proposition 8. The relier has an incentive to over-rely in order to force the producer to honor the contract.

Proposition 10:

Suppose that there exists an r such that $v(r) - r - p \geq 0$. Let $A^*(p,\alpha)$ be the values of r which maximize $a^*(\cdot,p,\alpha)$. Then

(i) $r \in A^*(p,a) \Rightarrow r \geq r_e$ for every (p,α).

(ii) If $v(r) - r - p$ is eventually negative for all large enough values of r, then $A^*(p,\alpha) \neq \emptyset$ for every (p,α).

(iii) Therefore under post-contract negotiations, reliance damages are less efficient than expectation damages.

Proof:

See Appendix. □

G. Behavior under Restitution Damages -- No Post-Contract Negotiations

Under restitution damages the seller pays expectation damages if he elects to sell to no one. However, if he elects to sell to a third party he pays the maximum of k - p and expectation damages. The seller, therefore, never has an incentive to sell to a third party. At best, he is indifferent between this option and performing. Since the buyer also receives the same return regardless, I will assume for analytical simplicity that the seller never sells to a third party. It is easy to see that the seller produces and honors the contract if and only if $c \leq v(r) - \overline{v}(r)$. Therefore the buyer's expected return to reliance can be written

$$e(r,p) = \int_{\{\theta \,:\, c(\theta) \leq v(r) - \overline{v}(r)\}} \max\begin{Bmatrix} v(r) \\ k(\theta) + \overline{v}(r) \end{Bmatrix} - r - p d\lambda$$

$$+ \int_{\{\theta \,:\, c(\theta) > v(r) - \overline{v}(r)\}} v(r) - r - p d\lambda \,. \tag{38}$$

Comparison of (38) with the expected return to reliance under expectation damages reveals that the reliance choice under restitution damages is less than or equal to the reliance choice under the other institution. This is because the buyer is only partially insured against the fact that it may be more efficient for a third party to receive the good. When $k > v(r) - \overline{v}(r)$ and $v(r) - \overline{v}(r) > c$, the buyer receives the good and resells it to the third party. Thus there is a chance that his reliance will have no return and the buyer takes this into account when he makes his reliance decision. Proposition 11 formally states this result.

Proposition 11:

The function $e(r,p)$ achieves its supremum and it does so independently of p. Let E be the set of all values of r which maximize e. Then

$$\sup E \leq r_e \tag{39}$$

Proof:

See Appendix. □

H. Behavior under Restitution Damages -- Post-Contract Negotiations

The breach inefficiency in restitution damages is that when $k > c > v(r) - \bar{v}(r)$, the seller does not produce even though it is efficient to do so. Therefore, we would expect negotiations to determine a side payment from the buyer to the seller and the seller to produce in this case. The expected return to reliance for the buyer is then

$$e^*(r,p,\alpha) = \int_{\{\theta\ :\ c(\theta) \leq v(r) - \bar{v}(r)\}} \max\left\{\begin{matrix} v(r) \\ k(\theta) - \bar{v}(r) \end{matrix}\right\} - r - p d\lambda$$

$$+ \int_{\{\theta\ :\ c(\theta) > v(r) - \bar{v}(r)\}} \left\{ v(r) - r - p + \alpha \max\{0, k(\theta) - c(\theta)\} \right\} d\lambda \tag{40}$$

This can be immediately rewritten as

$$e^*(r,p,\alpha) = e(r,p)$$

$$+ \alpha \int_{\{\theta \,:\, k(\theta) > c(\theta) > v(r) - \overline{v}(r)\}} k(\theta) - c(\theta)d\lambda. \qquad (41)$$

When α is zero the buyer expects to receive no rents from negotiations and he views the situation as identical to one where no post-contract negotiations occur. As α grows the buyer has less incentive to over-rely in order to force the seller to honor more often. This is because the buyer expects to do fairly well even when the seller breaches. Not surprisingly, therefore, the buyer's reliance decision grows smaller and more efficient as the value of α increases. Proposition 12 formally states this result.

Proposition 12:

The function $e^*(r,p,\alpha)$ achieves its supremum over r for every p and α, and this is done independently of p. Let $E^*(\alpha)$ denote the set of values for r which maximize $e^*(r,p,\alpha)$. Let $\alpha_1 < \alpha_2$, $r_1 \in E^*(\alpha_1)$, and $r_2 \in E^*(\alpha_2)$. Then

(i) $E^*(0) = E$.

(ii) $\sup E^*(\alpha_2) \leq \sup E^*(\alpha_1)$ and $\inf E^*(\alpha_2) \leq \inf E^*(\alpha_1)$.

(iii) A larger value of α results in at least as efficient an outcome. That is, $f(r_1) \leq f(r_2)$.

(iv) E is upper hemi continuous.

Proof:

See Appendix. □

I Behavior under Ideal Restitution Damages

Under ideal restitution damages the breaching seller pays the buyer

$$\max\{k - p, v(r) - \bar{v}(r) - p\}. \tag{42}$$

Ideal restitution damages thus protect the buyer's expectation under the assumption that the buyer would have sold to the third party if this was profitable. The buyer's return to reliance is thus

$$m(r,p) = \int_\Theta \max\{k(\theta) + \bar{v}(r) - r - p, v(r) - r - p\} d\lambda. \tag{43}$$

Figures 3 and 4 diagram the nature of the buyer's return to to reliance under both types of restitution damages.

FIGURE 3

BUYER'S RETURN TO RELIANCE UNDER RESTITUTION DAMAGES -- POST-CONTRACT NEGOTIATIONS

θ is such that the largest of $\{v - \bar{v}, c, k\}$ is	Value to Buyer
$v - \bar{v}$	$v - r - p$
c	$\max\left\{ \begin{matrix} v \\ k + \bar{v} \end{matrix} \right\} - r - p$
k, and $v - \bar{v} \geq c$	$k + \bar{v} - r - p$
and $c > v - \bar{v}$	$v - r - p + \alpha\ (k - c)$

FIGURE 4

BUYER'S RETURN TO RELIANCE UNDER IDEAL RESTITUTION DAMAGES

θ is such that the largest of $\{v - \bar{v}, c, k\}$ is	Value to Buyer
$v - \bar{v}$	$v - r - p$
c	$\max\left\{\begin{matrix} v \\ k + \bar{v} \end{matrix}\right\} - r - p$
k	$k + \bar{v} - r - p$

The only difference between the two is the returns when $k > c > v - \bar{v}$. In this case, reliance is valuable to the buyer operating under restitution damages because it increases his return in the absence of cooperation and thus increases his negotiation strength. Under ideal restitution damages the buyer is already assigned the property right to third party sales and there is no incentive for the buyer to attempt to increase his negotiation strength. The extra marginal return in the former case means that reliance is higher.

Recall that under expectation damages the buyer receives a return to his reliance even when the seller sells to a third party. This is because the buyer's damage award, $v(r) - \bar{v}(r) - p$, is an increasing function of r. However, there is no joint return to the buyer's reliance when the seller sells to a third party; the reliance is simply sold for scrap. This divergence between the private and joint evaluation of the marginal return to reliance contributes to the buyer's over-investment in reliance relative to the level which maximizes joint profits. Imposition of ideal restitution damages amounts to a full correction of this divergence between the marginal private and joint return to reliance. Under ideal restitution damages the buyer's damages depend on the price paid by the third party, $k(\theta)$, instead of $v(r)$. Since k does not depend on r, the buyer's marginal return to reliance now equals the marginal joint return. Therefore it is not surprising that ideal restitution damages result in a smaller and more efficient level of reliance than expectation damages. It is easy to see that ordinary restitution damages produce a partial correction for this divergence and that the correction increases with α. Therefore ordinary restitution damages produce reliance decisions midway between the other two, with the decision becoming closer to that of ideal restitution damages as α grows larger.

Recall that under expectation damages there is also a second contributor to an overly large reliance decision. In the case where cost rises dictate that the seller not produce at all, the buyer still receives a return to his reliance from the damage award. However, there is no joint return to reliance in this case. This distortion persists under ideal restitution damages, as is clear from Figure 4. Therefore, even ideal restitution damages produce a reliance decision larger than that which maximizes joint profits. Although it removes the distortion associated with third party offers, it does not affect the distortion associated with cost increases. This suggests, and it is in fact easy to show, that in a world where third party offers do not occur, ideal restitution, restitution and expectation damages are equivalent. The next two sections will show that specific performance begins where restitution damages leave off. Specific performance always produces a full correction of the third party offer distortion, and as α grows it also corrects progressively for the cost rise distortion so that when α equals 1, specific performance produces an efficient reliance decision. Proposition 13 summarizes ideal restitution damages' properties.

Proposition 13:

The function m(r,p) achieves its supremum independently of p. Let M be the values of r such that m achieves its supremum. Then

$$\inf \mathcal{R} \leq \inf \mathcal{M} \text{ and } \sup \mathcal{R} \leq \sup \mathcal{M} \tag{44}$$

$$\sup \mathcal{M} \leq \sup \mathcal{E}^*(1) \text{ and } \inf \mathcal{M} \leq \inf \mathcal{E}^*(1) \tag{45}$$

Proof:

See Appendix. □

J. Specific Performance -- No Post-Contract Negotiations

The damage remedy producing the most efficient reliance decision thus far, ideal restitution damages, is probably not implementable due to the court's inability to determine the value of k in the absence of a transaction occurring. Fortunately, specific performance will be seen to produce the identical incentives relating to the buyer's reliance decision as ideal restitution damages. In the next section it will be seen that specific performance induces an even more efficient reliance decision to the extent that the buyer believes he will capture rents from post-contract negotiations.

The buyer now always receives the good. His expected return to reliance is thus

$$s(r,p) = \int_{v(r) - \bar{v}(r) \geq k(\theta)} v(r) - r - p d\lambda$$

$$+ \int_{v(r) - \bar{v}(r) < k(\theta)} k(\theta) + \bar{v}(r) - r - p d\lambda \qquad (45)$$

Since he always receives the good his return is simply the maximum of the return he can receive by using it himself or selling to the third party. This is exactly what the seller receives under ideal restitution damages. That is, $s(r,p) = m(r,p)$. Therefore reliance choice under both institutions is the same. Let $\mathcal{S}$ be the maximizing choices of $s(r,p)$. We have proven

Proposition 14:

$$\mathcal{S} = \mathcal{M}$$

Proof:

As above. □

K. Specific Performance -- Post-Contract Negotiations

The inefficiency, given reliance, of specific performance is that the good is always produced, even when cost rises dictate that joint profits would be maximized by no production. Post-contract negotiations might be expected to resolve such a situation. The seller could offer the buyer a side payment in lieu of performance which would render them both better off. As usual, the buyer's expectations of the size of this side payment affect his expected return to reliance and thus his reliance decision. Once again α is the fraction of the increased joint profits that the buyer expects to receive. Figure 5 diagrams the buyer's expected return to reliance.

FIGURE 5

BUYER'S RETURN TO RELIANCE UNDER
SPECIFIC PERFORMANCE -- POST-CONTRACT NEGOTIATIONS

θ is such that the largest of $\{v - \overline{v}, c, k\}$ is	Value to Buyer
$v - \overline{v}$	$\max\{v - \overline{v}, k\} - r - p$
k	$\max\{v - \overline{v}, k\} - r - p$
c	$(1 - \alpha)[\max\{v - \overline{v}, k\} - r - p]$ $+ \alpha[c + \overline{v} - r - p]$

As α grows larger, the buyer's returns when costs increases make production unprofitable begin to depend more on the rents he can negotiate from the seller in exchange for allowing him out of the contract. Mathematically, this means that his returns in this case begin to depend less and less on his own reliance choice and more on the size of the seller's cost over-run. From Figure 1 it is clear that this is also the case for the joint return to reliance. In fact, when α equals 1, it is clear from comparing the two figures that the returns only differ by a constant, $p - c(\theta)$. Therefore the same choices of r maximize the aggregate return as the buyer's return under specific performance when α equals 1. That is, the buyer's choice of reliance maximizes joint profits in this case.

As explained in Section I, expectation damages produces an over-investment in reliance because the buyer receives a return to his reliance when it is efficient for a third party to receive the good or for the good not to be produced at all even though the buyer's reliance is not used in these cases and therefore does not increase joint profits. Therefore, *ex ante*, the buyer over-values reliance and over-invests in it.

Ideal restitution damages and specific performance when α is zero amount to a complete correction of the divergence between the private and joint returns to reliance for the case where the efficient course of action is for the third party to receive the good. As α grows larger, specific performance progressively corrects for the other distortion as well, so that when α equals 1, specific performance produces a perfectly efficient reliance decision. The buyer makes a more efficient decision as α grows because he expects to receive more and more of the gains from an efficient decision on his part.

The buyer's return to reliance in this case is

$$s^*(r,p,\alpha) = \int_{P_1^*(r)} v(r) - r - pd\lambda + \int_{P_2^*(r)} k(\theta) + \bar{v}(r) - r - pd\lambda$$

$$+ \int_{P_3^*(r)} \left\{\bar{v}(r) - r - p + \alpha c(\theta) + (1 - \alpha) \max\{k(\theta), v(r) - \bar{v}(r)\}\right\} d\lambda. \quad (46)$$

Proposition 15 summarizes the properties of s^*.

Proposition 15:

The function $s^*(\cdot,p,\alpha)$ achieves its supremum independently of p. Let $S^*(\alpha)$ denote the set of values which maximize $s^*(\cdot,p,\alpha)$. Let $\alpha_1 < \alpha_2$. Then

(i) $S^*(0) = S$.

(ii) $S^*(1) = R$.

(iii) $\inf S^*(\alpha_2) \leq \inf S^*(\alpha_1)$ and $\sup S^*(\alpha_2) \leq \sup S^*(\alpha_1)$.

(iv) Higher values of α produce more efficient outcomes. That is, if $r_i \in S^*(\alpha_i)$, then $f(r_2) \geq f(r_1)$.

(v) S^* is upper hemi continuous.

Proof:

See Appendix. □

L. Liquidated Damages

The optimality of specific performance depends on post-contract negotiations successfully occurring and for the buyer to believe he will be successful in capturing most of the rents up for negotiation. Furthermore, we have not considered the transactions costs of these negotiations or the effects of the *ex ante* increase in risk created by relying on the outcome of post-contract negotiations. For these reasons it may well be that damage institutions not relying on post-contract negotiations are generally superior to those that do. It is easy, for example, to create examples where agents averse to the risk of post-contract negotiations prefer expectation damages to specific performance.

Proposition 16:

Agents averse to the risk of negotiations may prefer expectation damages to specific performance.

Proof:

See Appendix. □

Fortunately, a damage measure does exist which induces efficient reliance on the part of the buyer but does not require post-contract negotiations. The buyer and seller can insert a value of lump-sum damages to be paid by the seller to the buyer in the event of breach which results in an efficient reliance choice by the buyer and an efficient breach decision by the seller.

Proposition 17:

Let r^* be any element of $\mathcal{R}$. Then a liquidated damages award of $v(r^*) - \bar{v}(r^*) - p$ induces a reliance choice in $\mathcal{R}$ for the buyer and an efficient breach choice for the seller.

Proof:

See Appendix. □

The intuition is clear. Expectation damages, $v(r) - \bar{v}(r) - p$, produce efficient breach behavior at r. As long as the buyer chooses r*, the seller thus exhibits efficient breach behavior. The buyer now chooses an efficient level of reliance because the damages he receives in the event of breach do not depend on his own level of reliance.

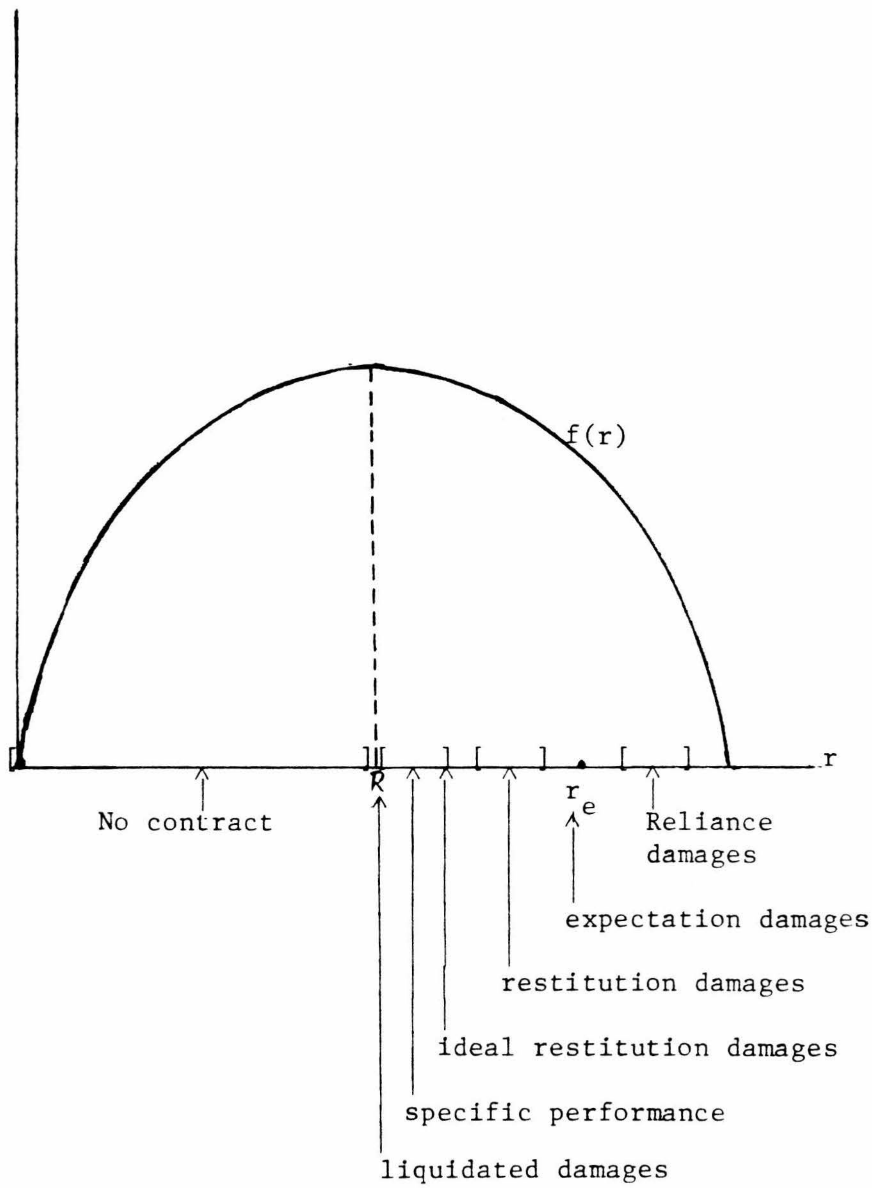
Since the buyer and seller can maximize their *ex ante* joint return by choosing liquidated damages of $v(r^*) - \bar{v}(r^*) - p$, presumably they would do so. This raises the question of why all contracts do not incorporate liquidated damages. First, the amount of reliance at stake may not always be large enough to be significant. Second, courts do not always enforce liquidated damages clauses (Goetz and Scott, 1977). Third, and possibly most important, however, v may be random. That is, v may be a function of θ as well. In this case, it is easy to prove that liquidated damages of $v(r^*,\theta) - \bar{v}(r^*,\theta) - p$ induce efficient behavior, but now contracts must specify an entire function instead of one number. Furthermore, various moral hazard problems are raised if θ is not observable. Therefore, as v becomes more variable, liquidated damages becomes a less satisfactory damage measure. This theory predicts that in cases where the cost of inefficient reliance is significant and v is not too variable, liquidated damages would be used.

M. Summary

Figure 6 provides a graphical summary of the results of the case where no market for substitute performance exists.

FIGURE 6

SUMMARY -- NO MARKET FOR SUBSTITUTE PERFORMANCE



Four points should be noted. First, the intervals for some damage measures occur because post-contract negotiations occur to resolve inefficient breach behavior at the given level of reliance. Each point in the interval corresponds to the buyer's reliance choice under a different value for his subjective negotiation strength, α. For no contract, specific performance, and restitution damages, the reliance choice moves towards R and becomes more efficient as α grows. Second, because the intervals are all disjoint, we get an unambiguous ranking of the institutions even if a different value for α is used for different institutions. Third, some of the endpoints of intervals may be shared and some intervals may collapse in degenerate situations. For example, when there are no third party offers the entire restitution damages interval collapses into r_e. Fourth, f is drawn as concave in Figure 6. In fact, f is not necessarily concave. The original work in this area (Shavell 1978) assumed that v was concave and then asserted this implied f was also concave. This assertion was false. All the results we would expect were f concave are instead proven on the basis of dividing the objective functions into non-decreasing and non-increasing parts. See Lemmas 3, 4, and 5 in particular. Since all the theorems that would be true were f concave are true, it is convenient to draw f as being concave.

V. OTHER CONSIDERATIONS

A number of factors were not considered which may affect the nature of the conclusions. The purpose of this section is to take note of some of these and speculate on their effects, suggesting directions for future research.

First, the different damage measures allocate risk differently. For example, restitution damages place the risk of third party offers on the buyer; expectation damages place this risk on the seller. Differential attitudes towards risk might therefore affect choice of a damage measure.

Second, the buyer and seller may possess different subjective evaluations of the likelihood of various outcomes of θ. Therefore, for example, the agent who thought a particular risk had less variance might be more willing to bear it. An important example concerns the case where the buyer has very little knowledge of θ at all, but the seller has good information on θ and knows he will in all likelihood honor the contract. To the seller there is little cost in agreeing to specific performance; to the buyer there is a large benefit, especially if the buyer would have to incur search costs in the event of breach. (Search costs are typically not included in expectation damages.) In some sense, by agreeing to specific performance the seller is communicating to the buyer that breach is unlikely.

Third, it was assumed that both parties could equally well resell the good to the third party. It may well be, however, that the seller has an advantage in selling to third parties because of fixed costs such as a showroom, repair services, reputation, etc. Therefore damage measures which transfer the property right to third party sales to the buyer (specific performance, ideal restitution damages, and restitution damages) might well involve a post-contract negotiation transferring this property right back to the seller in the event of a large third party offer. Post-contract negotiations result in extra risk and transactions costs.

Fourth, search may be required to ferret out third party offers. If, as seems likely, the seller is the efficient searcher, transfer of the property right to third party sales to the buyer would result in too little search occurring.

This transfer of the property right to third party offers has been taken note of in the literature on specific performance (Kronman 1978; Schwartz 1979). Kronman attempted to argue that specific performance would be typically more desired for unique goods on the basis of the first point above. Schwartz pointed out that Kronman's speculations were not correct. It is true that specific performance is equivalent to expectation damages when a market for substitute performance exists, and is thus not necessary. Within the class of goods for which no market for substitute performance exists, the second point above suggests a type of good for which specific performance might be desirable. However, the third and fourth points above suggest types of goods for which specific performance would be undesirable. Therefore, although the absence of a market for substitute performance is necessary for specific performance to be useful, it is by no means sufficient.

Fifth, when damage measures produce inefficient breach behavior at the given reliance level and depend on post-contract negotiations to resolve this, transactions costs or aversion to the risk of negotiations may produce an advantage for damage measures which do induce efficient breach behavior. See Proposition 16 and the discussion surrounding it.

A sixth, related point is that the choice of an efficient level of reliance requires knowledge of $k(\theta)$ and $c(\theta)$. The damage measures which outperform expectation damages require the buyer to possess progressively more information as the reliance decision becomes progressively more efficient. For example, under restitution damages the buyer must know $k(\theta)$ to calculate his optimal level of reliance; under specific performance he must know $k(\theta)$ and $c(\theta)$. Therefore an implicit assumption in the conclusion that various measures outperform expectation damages is that the buyer has the information to make an optimal reliance choice. In a situation where the buyer has no knowledge of $c(\theta)$ or $k(\theta)$, the institution of expectation damages which requires that the buyer only know his own private information, $v(r)$, may produce the same reliance decision as institutions which outperform it under conditions of fuller information. Points five and six may help explain the prevalent use of expectation damages in the courts. The buyer often may not possess sufficient information for other institutions to produce a better reliance decision than expectation damages and expectation damages results in no post-contract negotiations.

Seventh, much less litigation may be involved in enforcing some damage measures than others. When the good or service contracted for is difficult to monitor, specific performance may be very expensive to enforce. Often, expensive litigation will be required to determine the value of the award under expectation damages.

Possibly liquidated damages involve the smaller enforcement cost in the case where they do not depend on difficult to measure factors (i.e., where v is not a function of a difficult to monitor θ).

Eighth, the courts often have difficulty in awarding "idiosyncratic" expectation damages. Even though a particular good or service may be worth well above the average value of such goods or services to a particular breachee, it is difficult to establish this and consequently expectation damages do not protect this "idiosyncratic" value. Liquidated damages or specific performance can do this. This point has been made persuasively in the literature (Goetz and Scott 1977).

VI. CONCLUSION

When a market for substitute performance exists, most of the problems considered in this paper vanish. The relier needs no assurances of performance to rely efficiently. The institutions of expectation damages and specific performance produce no moral hazard and result in efficient allocation of resources. That is, society finds itself in the unfortunate position that the existence of a problem for the institution of contracts to solve is also sufficient to cause contracts to create their own inefficiency. This paper is an analysis of the capabilities of various contract institutions for coping with such a perverse market failure.

The absence of a competitive market means that a party who engages in reliance in anticipation of a future exchange worsens his negotiating position in the future negotiation over price of exchange. This results in less than efficient levels of reliance. Contracts "solve" this problem, but only at the cost of providing that the relier receive some private return from reliance even in cases where some exogenous happening renders exchange inefficient and the reliance has no social return. As with many other types of insurance, a moral hazard is created. In this case the relier tends to over-rely.

Reliance damages result in the least efficient allocation of goods regardless of whether or not post-contract negotiations occur and regardless of the level of confidence the relier has concerning this negotiating strength. Restitution damages tend to produce a more efficient reliance choice than expectation damages because under the former the relier takes into account the fact that sales to a third party may make his reliance useless. Assuming that post-contract negotiations occur, specific performance is even more efficient. If the relier has absolutely no confidence in his negotiating strength, his reliance choice is largest. As his confidence grows he takes more and more account of the fact that rises in cost of production may render performance inefficient and thus he relies more efficiently. In the limit, as he expects to receive all the rent from any negotiaiton, his reliance decision is efficient. The buyer and seller can choose a level of liquidated damages, which induces efficient behavior.

The formal analysis of this paper highlights the crucial role of information in the process of contracting. Essentially, it seems that although other institutions may produce more efficient behavior than expectation damages, the institution currently being used, they require more information and/or negotiations. Under conditions of poor information, agents averse to the variance generated by poor information might well find that expectation damages are the most efficient contract institution.

Current law frowns upon specific performance and liquidated damages clauses, often not enforcing such agreements (Schwartz 1979; Goetz and Scott 1977). The economists' simple-minded statement that freedom of choice can only always make everyone better off is not sufficient to sway the courts. They claim that the opportunity to insert such clauses in contracts often allows more exercise of monopoly power and other unfair bargaining power. This may well be and is worthy of analysis in itself. Another line of counterargument is to identify specific reasons why parties to a contract might find it to their mutual advantage to insert such clauses into their contracts. This tactic has been followed by previous authors (Kronman 1978; Schwartz 1979; Goetz and Scott 1977). This paper adds another reason why parties to a contract may prefer specific performance or liquidated damages over more conventional remedies.

APPENDIX

To prove Proposition 2, it is easiest to first establish two simple lemmas.

Lemma 1:

Let S be a set. Let h_i be a real valued function defined on R x S for i = 1,...,n. Define h: R x S → R by

$$h(x,s) = \max_{i \in \{1,\ldots,n\}} \{h_i(x,s)\}. \quad (A-1)$$

Suppose each h_i is continuous in its first variable uniformly over its second variable. That is, for every $i \in \{1,\ldots,n\}$, $x_o \in R$ and $\varepsilon > 0$ there exists a $\delta > 0$ such that for every $s \in S$ and $x \in R$

$$|x - x_o| < \delta \Rightarrow |h_i(x,s) - h_i(x_o,s)| < \varepsilon. \quad (A-2)$$

Then h is also continuous in its first variable uniformly over its second variable.

Proof:

Choose any $x_o \in R$ and $\varepsilon > 0$. Then there exists a $\delta > 0$ such that for every $i \in \{1,\ldots,n\}$, $s \in S$ and $x \in R$

$$|x - x_o| < \delta \Rightarrow |h_i(x,s) - h_i(x_o,s)| < \varepsilon. \quad (A-3)$$

Now let $i^*(x)$ and i^{**} be the elements of $\{1,\ldots,n\}$ which, respectively, maximize $h_i(x,s)$ and $h_i(x_o,s)$. Then

$$|h(x,s) - h(x_o,s)| = |h_{i*(x)}(x,s) - h_{i**}(x_o s)| \qquad \text{(A-4)}$$

First suppose that $h_{i*(x)}(x,s) - h_{i**}(x_o,s) \geq 0$. Then it is less than or equal to $h_{i*(x)}(x,s) - h_{i*(x)}(x_o,s)$ which is less than ε by (A-3). Second, suppose that $h_{i*(x)}(x,s) - h_{i**}(x_o,s) < 0$. Then it is less than or equal to $h_{i**}(x,s) - h_{i**}(x_o,s)$ in absolute value which is less than ε in absolute value by (A-3). □

Lemma 2:

Let S be a measurable bounded subset of R^n. Let h be a real valued function defined over R x S which is continuous in its first variable uniformly over its second variable and measurable in its second variable. Define $g:R \to R$ by

$$g(x) = \int_S h(x,s)ds. \qquad \text{(A-5)}$$

Then g is continuous.

Proof:

Choose any $x_o \in R$ and $\varepsilon > 0$. Then select a δ such that

$$|x - x_o| < \delta \Rightarrow |h(x,s) - h(x_o,s)| < \frac{\varepsilon}{\lambda(S)}. \qquad \text{(A-6)}$$

Then

$$|g(x) - g(x_o)| \leq \int_S \frac{\varepsilon}{\lambda(S)} d\lambda = \varepsilon. \tag{A-7}$$

□

Proposition 2

It is clear that f can be rewritten as

$$f(r) = \int_{[0,1]^2} \max \left\{ \begin{array}{l} v(r) - r - c(\theta) \\ \overline{v}(r) - r + k(\theta) - c(\theta) \\ v(r) - r - k(\theta) \\ \overline{v}(r) - r \end{array} \right\} d\lambda. \tag{A-8}$$

Each of the four functions in brackets is continuous in r uniformly over θ. Therefore by Lemma 1, so is their maximum. Therefore by Lemma 2 f is continuous.

The function $\overline{v}(r) - r$ is always decreasing in r. Therefore, for $r > r_e$ all four functions are larger at r_e than r. Therefore so is f. Therefore f achieves its supremum on $[0,r_e]$. □

Proposition 3

It is clear that the buyer and seller act so that $P_i = P_i^*(r)$ for i = 1,2,3,4. Therefore the buyer's expected return as a function of r is

$$b(r) = \int_{P_1^*(r) \cup P_3^*(r)} v(r) - r - k(\theta) d\lambda + \int_{P_2^*(r) \cup P_4^*(r)} \overline{v}(r) - r d\lambda. \tag{A-9}$$

Now rewrite b(r) as

$$b(r) = f(r) + \int_{P_1^*(r) \cup P_2^*(r)} c(\theta) - k(\theta) d\lambda. \tag{A-10}$$

However, $P_1^*(r) \cup P_2^*(r)$ equals

$$\{(\theta): c(\theta) \leq k(\theta)\} \tag{A-11}$$

which does not depend on r. Therefore b(r) and f(r) differ only by a constant, and they have the same maximizing values. □

Proposition 4

It is clear that the buyer and seller will act so that $P_i = P_i^*(r)$ for i = 1,2,3,4. Therefore the buyer's expected return as a function of r and p, the contract price, is

$$e(r,p) = \int_{P_1^*(r) \cup P_3^*(r)} v(r) - r - p d\lambda + \int_{P_2^*(r) \cup P_4^*(r)} \bar{v}(r) - r + k(\theta) - p d\lambda. \tag{A-12}$$

Now rewrite e(r,p) as

$$e(r,p) = f(r) + \int_{P_1^*(r) \cup P_2^*(r)} c(\theta) - p d\lambda + \int_{P_3^*(r) \cup P_4^*(r)} k(\theta) - p d\lambda. \tag{A-13}$$

However, as in Proposition 3, $P_1^*(r) \cup P_2^*(r)$ does not vary with r. Neither does $P_3^*(r) \cup P_4^*(r)$. Therefore e(r,p) differs from f(r) by a constant. □

Lemma 3 is used to prove Proposition 7.

Lemma 3:

Let h and g be real valued functions defined on R. Let g be decreasing. Choose α_1 and α_2 in (0,1]. Let $\alpha_1 > \alpha_2$. Suppose that r_i maximizes

$$h(r) + \frac{(1 - \alpha_i)}{\alpha_i} g(r) \tag{A-14}$$

for i = 1,2. Then $r_1 \geq r_2$. Furthermore, if $\alpha_i \neq 1$, then $r < r_i$ implies that $h(r) < h(r_i)$.

Proof:

Suppose that $r_2 > r_1$. Then $g(r_2) < g(r_1)$. As well $\frac{(1 - \alpha_1)}{\alpha_1} < \frac{(1 - \alpha_2)}{\alpha_2}$. Therefore we have

$$\left[\frac{(1 - \alpha_2)}{\alpha_2} - \frac{(1 - \alpha_1)}{\alpha_1}\right] g(r_2) < \left[\frac{(1 - \alpha_2)}{\alpha_2} - \frac{(1 - \alpha_1)}{\alpha_1}\right] g(r_1) \tag{A-15}$$

As well since r_1 is a maximizing value of (A-14) for α_1,

$$h(r_2) + \frac{(1 - \alpha_1)}{\alpha_1} g(r_2) \leq h(r_1) + \frac{(1 - \alpha_1)}{\alpha_1} g(r_1) \tag{A-16}$$

Add (A-15) and (A-16) to yield

$$h(r_2) + \frac{(1 - \alpha_2)}{\alpha_2} g(r_2) < h(r_1) + \frac{(1 - \alpha_2)}{\alpha_2} g(r_1) \tag{A-17}$$

which contradicts the definition of r_2.

Therefore r_2 must be less than or equal to r_1.

For the last part, if $r < r_i$, then the fact that g is decreasing implies $h(r) < h(r_i)$. □

Proposition 7

Rewrite $b(r,\alpha)$ as

$$b(r,\alpha) = \alpha\left(f(r) + \int_{\{\theta\ :\ c\ <\ k\}} c(\theta) - k(\theta)d\lambda\right) + (1 - \alpha)(\bar{v}(r) - r) \qquad \text{(A-18)}$$

by performing the required algebra. If $\alpha = 0$, $N(\alpha)$ clearly equals $\{0\}$. Therefore for any $\alpha \neq 0$, sup $N(0) \leq$ inf $N(\alpha)$. Now consider the case where $\alpha \neq 0$. Rewrite b as

$$b(r,\alpha) = \alpha\left[f(r) + \frac{(1 - \alpha)}{\alpha}(\bar{v}(r) - r) + \int_{\{\theta\ :\ c\ <\ k\}} c(\theta) - k(\theta)d\lambda\right] \qquad \text{(A-19)}$$

Since $r > r_e$ implies $f(r) < f(r_e)$, $\bar{v}(r) - r$ is decreasing, and since both $f(r)$ and $\bar{v}(r) - r$ are continuous, $b(r,\alpha)$ achieves its supremum Lemma 3 now yields results (iii) and (iv). Result (v) follows directly from a theorem in Debreu (1959), page 1.9. □

Proposition 8

Rewrite equation (34) as

$$a(r,p) = (v(r) - r - p)\ \lambda(H(r,p)). \qquad \text{(A-20)}$$

For $r < r_e$, $v(r) - r - p < v(r_e) - r_e - p$. As well, $\lambda \cdot H(\cdot,p)$ is nondecreasing in r. Therefore, all of the global maxima to $a(\cdot,p)$ occur in $[r_e,\infty)$.

For part (ii) it is sufficient to show that $\lambda \cdot H$ is upper semicontinuous in r. This is true because $\lambda \cdot H(r,p)$ is actually $F(p - v(r) + r, p - v(r) + r)$ where F is the distribution function of the random variables c and k. The function F is by definition nondecreasing and continuous from the right so is upper semicontinuous. □

Proposition 1C

Rewrite $a^*(r,p,\alpha)$ as follows

$$a^*(r,p,\alpha) = \alpha \int_{P_1^*(r)} v(r) - \max\begin{Bmatrix} r + p \\ k(\theta) + \bar{v}(r) \\ c(\theta) + \bar{v}(r) \end{Bmatrix} d\lambda$$

$$+ (1 - \alpha) \int_{H(r,p)} v(r) - r - p d\lambda. \tag{A-21}$$

As for Proposition 8, both integrals are larger at r_e than at any $r < r_e$. Therefore this is also true for any convex combination of the two integrals. Consequently $a^*(\cdot,p,\alpha)$ is maximized on $[r_e,\infty)$ if at all.

For existence, the second integral is an upper semicontinuous function of r as proved in Proposition 8. The first integral can be rewritten as

$$\int_{[0,1]} \max \begin{Bmatrix} v(r) \\ k(\theta) - \overline{v}(r) \\ c(\theta) + \overline{v}(r) \end{Bmatrix} - \max \begin{Bmatrix} r + p \\ k(\theta) + \overline{v}(r) \\ c(\theta) + \overline{v}(r) \end{Bmatrix} d\lambda. \qquad \text{(A-22)}$$

Now in much the same fashion as for Proposition 2, it can be shown that (A-22) is a continuous function of r. Therefore a* is an upper semicontinuous function of r for every (p,α). If $v(r) - r - p$ is eventually negative for large enough r, it is clear that $a^*(r,p,\alpha)$ will also be negative for large enough r. As a consequence, it must achieve its supremum somewhere. □

Proposition 11

Expression (39) is proven in the same fashion as the previous proofs, by dividing e into the function whose supremum occurs at tye bound and a function which is nondecreasing. Once (39) is proven, for existence it is sufficient to show that e is upper semicontinuous. This is done by the same type of argument as in Propositions 2 and 8. □

Lemma 4:

Let h and g be real valued functions defined on $\mathcal{R}$. Let g be nonincreasing. Choose α_1 and α_2 in [0,1] with $\alpha_1 < \alpha_2$. Suppose that $\mathcal{R}_i$ is the set of values for r which maximize

$$h(r) + \alpha_i g(r). \qquad \text{(A-23)}$$

Let $r_i \varepsilon R_i$. Then

(i) $\sup R_2 \leq \sup R_1$

(ii) $\inf R_2 \leq \inf R_1$

(iii) $r < r_i$ implies $h(r) \leq h(r_i)$.

Proof:

(i) Suppose that there exists an $r_2 \varepsilon R_2$ such that $r_2 > \sup R_1$. Then consider any $r_1 \varepsilon R_1$. We know that

$$h(r_1) + \alpha_1 g(r_1) > h(r_2) + \alpha_1 g(r_2) \tag{A-24}$$

because $r_2 > \sup R_1$.
As well,

$$(\alpha_2 - \alpha_1) g(r_1) \geq (\alpha_2 - \alpha_1) g(r_2) \tag{A-25}$$

because g is in increasing and $r_2 > \sup R_1$. Adding (A-24) and (A-25) yields

$$h(r_1) + \alpha_2 g(r_1) > h(r_2) + \alpha_2 g(r_2), \tag{A-26}$$

contradicting the definition of r_2. Therefore, for every $r_2 \varepsilon R_2$, $r_2 \leq \sup R_1$, or equivalently, $\sup R_2 \leq \sup R_1$.

(ii) This is proven in a similar fashion to (i).

(iii) The fact that g is non-increasing means

$$-\alpha_i g(r) \leq - \alpha_i g(r_i). \tag{A-27}$$

Also, since $r_i \in R_i$,

$$h(r) + \alpha_i g(r) \leq h(r_i) + \alpha_i g(r_i). \tag{A-28}$$

Addition of (A-27) and (A-28) yields

$$h(r) \leq h(r_i). \tag{A-29}$$

□

Proposition 12:

Part (i) is obvious. Parts (ii) and (iii) follow immediately from Lemma 4. Existence of a maximum follows as usual from upper semicontinuity. □

Proposition 13:

Rewrite (43) as

$$m(r,p) = e^*(r,1) + \int_{\{\theta \,:\, k > c > v - \bar{v}\}} \{c(\theta) - (v(r) - \bar{v}(r))\} d\lambda \tag{A-30}$$

The above integrand is decreasing in r and non-negative. The region being integrated over shrinks as r increases. Therefore the integral in (A-30) is a non-increasing function of r. From this, (45) is clear.

Now rewrite (43) as

$$m(r,p) = f(r) + c(\theta) - p + \int_{\{\theta \,:\, c(\theta) > k(\theta) \wedge c(\theta) > v(r) - \bar{v}(r)\}} \{\max(k(\theta), v(r) - \bar{v}(r)) - c(\theta)\} d\lambda \tag{A-31}$$

The integrand is negative and increases in r. The domain of integration decreases with r. Therefore, the integral is a non-decreasing function of r. This proves (44).

Existence and upper hemi-continuity is proved by the usual upper semicontinuity argument as in Proposition 7. □

Lemma 5:

Let h and g be real valued functions defined on $\mathcal{R}$. Let g be non-decreasing. Choose α_1 and α_2 in [0,1] with $\alpha_1 < \alpha_2$. Suppose that $\mathcal{R}_i$ is the set of values for r which maximize

$$h(r) + \alpha_i g(r). \tag{A-32}$$

Let $r_i \in \mathcal{R}_i$. Then

(i) $\inf \mathcal{R}_1 \leq \inf \mathcal{R}_2$

(ii) $\sup \mathcal{R}_1 \leq \sup \mathcal{R}_2$

(iii) $r > r_i$ implies $h(r) \leq h(r_i)$.

Proof:

The proof is substantially the same as that of Lemma 4. □

Proposition 15:

First rewrite $s^*(r,p,\alpha)$ as

$$s^*(r,p,\alpha) = v(r) - r - p + \int_{P_2^*(r)} \bar{v}(r) - v(r) + k(\theta)d\lambda$$

$$+ \int_{P_3^*(r)} \bar{v}(r) - v(r) + c(\theta) + (1 - \alpha)\max\{k(\theta), v(r) - \bar{v}(r)\}d\lambda. \tag{A-33}$$

It is easy to see that both integrals are non-increasing functions of r. Therefore if s^* achieves its supremum, it does so on $[0,r_e]$. The usual proof of upper semicontinuity thus establishes existence and upper hemi-continuity of S^*.

Now rewrite $s^*(r,p,\alpha)$ as

$$s^*(r,p,\alpha) = f(r) + \int_{[0,1]} c(\theta) - pd\lambda$$

$$+ (1 - \alpha)\int_{P_3^*(r)} \max\{k(\theta), v(r) - \bar{v}(r)\} - c(\theta)\ d\lambda. \qquad \text{(A-34)}$$

It is easy to see that the second integral is a non-decreasing function of r. □

Proposition 16:

The proof of this proposition is a counterexample. Suppose that the buyer is averse to the risks of negotiations. Model this by assuming that although the buyer believes the expected value of his negotiating strength is such that α is α^*, he uses $\alpha = \alpha_b$ where α_b is less than α^* to evaluate the desirability of any such gamble. This is a type of certainty equivalence. In the limit the buyer may have no basis at all for forming an expected value. He might then exhibit maxi-min behavior, attempting to guarantee himself some minimum level of income even in the worst case. For the buyer this means choosing $\alpha_b = 0$. To the extent that the buyer is averse to risks associated with the negotiation process, he will choose α_b less than α^*, the true value of α.

I assume that the seller is risk neutral and that there are no third party offers. More general counterexamples can, of course, be created, but since this is essentially a counterexample to the previous results, I select the simplest case.

This example shows that if the buyer is risk averse enough, expectation damages are actually more efficient than specific performance. The intuition behind this result is straightforward. As the buyer becomes more risk averse, his reliance choice under specific performance becomes less and less superior to the reliance choice under expectation damages. That is, the advantage to specific performance declines. This fact alone would cause specific performance to be equally efficient to expectation damages in the limit. However, another factor is also at work, besides the reliance choice becoming less efficient. As well, the buyer begins to discount the expected return from negotiations. At some point these factors invariably combine to produce a larger efficiency gain for expectation damages. The formal statement is as follows.

Formal Statement of Counterexample:

Suppose that the buyer compares returns from specific performance to those from expectation damages by using $\alpha = \alpha_b$. Let the actual value of α be α^*. Suppose that there is a positive probability of cost overruns. That is,

$$\int_{P_3^*(r_e)} c(\theta)d\lambda > 0.$$

Then

(i) For every $\alpha^* > 0$ there exists an $\alpha_b^* > 0$ such that

$\alpha_b < \alpha_b^*$ implies that expectation damages are more efficient than specific performance.

(ii) If $\alpha^* = 0$, then $\alpha_b = 0$ implies that expectation damages and specific performance are equally efficient.

<u>Proof</u>:

The seller evaluates the two returns by comparing their expected values. The buyer evaluates the two returns by comparing the certain return from expectation damages to the discounted expected return from specific performance. It is clear in this case that one institution is more efficient than the other if and only if the sum of the buyer's and seller's evaluations is greater. (A side payment is always possible such that the institution producing the greatest aggregate value will result in both parties being better off than under the second institution.) Therefore it is sufficient to show that the sum of evaluations from expectation damages is less than that from specific performance. The former number is $f(r_e)$. The latter is

$$S^*(r_s,p,\alpha_b) + f(r_s) - S^*(r_s,p,\alpha^*) \qquad \text{(A-35)}$$

where r_s is the element of $S^*(\alpha_b)$ chosen by the buyer. Let $G(\alpha_b)$ be the set of all possible values for (A-35).

$$G(\alpha_b) = \{S^*(r_s,p,\alpha_b) + f(r_s) - S^*(r_s,p,\alpha^*) : r_s \in S^*(\alpha_b)\}. \qquad \text{(A-36)}$$

Let $g(\alpha_b)$ be the supremum of $G(\alpha_b)$.

$$g(\alpha_b) = \sup G(\alpha_b). \qquad \text{(A-37)}$$

It is easy to observe that $g(0)$ equals $f(r_e)$ when $\alpha^* = 0$ and is less than $f(r_e)$ when $\alpha^* > 0$. Therefore (ii) is proven. To prove (i) we must show that g is upper semicontinuous. This follows from a theorem in Debreu (Debreu 1959), page 1.9. □

Proposition 17:

First consider the seller. The seller receives $p - c(\theta)$ if he performs, $k(\theta) - c(\theta) - v(r^*) + \overline{v}\,(r^*) + p$ if he sells to a third party, and $-v(r^*) + \overline{v}(r^*) + p$ if he breaches. The seller's breach decision is therefore as follows:

θ is such that the largest of $\{v(r^*) - \overline{v}(r^*), k(\theta), c(\theta)\}$ is	The seller
$v(r^*) - \overline{v}(r^*)$	honors contract
$k(\theta)$	sells to third party
$c(\theta)$	does not produce

Now consider the buyer. Let d be the damages. Then the buyer's return to reliance is

$$\ell(r,p) = \int_{P_1^*(r^*)} v(r) - r - p \, d\lambda + \int_{P_2^*(r^*) \,\cup\, P_3^*(r^*)} d + \overline{v}(r) - r \, d\lambda. \qquad \text{(A-38)}$$

Rewrite this as

$$\ell(r,p) = F(r,P_1^*(r^*),P_2^*(r^*),P_3^*(r^*)) + \int_{P_1^*(r^*)} c(\theta) - pd\lambda + \int_{P_2^*(r^*)} d + c(\theta) - k(\theta)d\lambda + \int_{P_3^*(r^*)} d\ d\lambda. \tag{A-39}$$

None of the integrals are functions of r. Therefore ℓ and F differ only by a constant and are maximized by the same values of r. To see that ℓ and F are maximized at r, suppose they are not, for contradiction. Then there exists an r' such that

$$F(r',P_1^*(r^*),P_2^*(r^*),P_3^*(r^*)) > F(r^*,P_1^*(r^*),P_2^*(r^*),P_3^*(r^*)). \tag{A-40}$$

By definition,

$$F(r',P_1^*(r'),P_2^*(r'),P_3^*(r')) \geq F(r',P_1^*(r^*),P_2^*(r^*),P_3^*(r^*)). \tag{A-41}$$

Combining (A-40) and (A-41) yields

$$F(r',P_1^*(r'),P_2^*(r'),P_3^*(r')) > F(r^*,P_1^*(r^*),P_2^*(r^*),P_3^*(r^*)). \tag{A-42}$$

Therefore any value of r chosen by the buyer results in the maximization of joint returns. □

REFERENCES

Barton, John H. "The Economic Basis of Damages for Breach of Contract." Journal of Legal Studies 1 (1972):277-304.

Birmingham, Robert L. "Legal and Moral Duty in Game Theory: Common Law Contract and Chinese Analogies." Buffalo Law Review 99 (1969):103-110.

Debreu, Gerard. Theory of Value, Cowles Foundation Monograph Number 17, New Haven and London: Yale University Press, 1959.

Diamond, P. and Maskin, E. "An Equilibrium Analysis of Search and Breach of Contract, I: Steady States." Bell Journal 10 (1979):282-316.

__________. and __________. "An Equilibrium Analysis of Search and Breach of Contract, II: A Non-Steady State Example." MIT Department of Economics Working Paper No. 237, April 1979.

Farnsworth, E. Allan. "Legal Remedies for Breach of Contract." Columbia Law Review 70 (1970):1145-1216.

Goetz, Charles J. and Scott, Robert E. "Liquidated Damages, Penalties and the Just Compensation Principle: Some Notes on an Enforcement Model and a Theory of Efficient Breach." Columbia Law Review 77 (1977):554-594.

Holmström, B. "Moral Hazard and Observability." Bell Journal of Economics 10 (spring 1979):74-91.

Joskow, Paul L. "Commercial Impossibility, the Uranium Market and the Westinghouse Case." Journal of Legal Studies 6 (1976): 119-176.

Klein, Benjamin; Crawford, Robert G.; and Alchian, Armen A. "Vertical Integration, Appropriable Rents, and the Competitive Contracting Process." Journal of Law and Economics 21 (1978):297-326.

Kronman, Anthony T. "Specific Performance." University of Chicago Law Review 45 (1978):351-382.

__________. and Posner, Richard A., eds., The Economics of Contract Law, Boston: Little, Brown and Company, 1979.

Leff, Arthur Allen "Injury, Ignorance and Spite -- The Dynamics of Coercive Collection." Yale Law Journal 80 (1970):1-29.

Landa, Janet and Grofman, Bernard. "Games of Breach and the Role of Contract Law in Protecting the Expectation Interest." University of California, Irvine: School of Social Sciences Research Report No. 34, June 1979.

Perloff, Jeffrey M. "Breach of Contract and the Foreseeability Doctrine of Hadley v. Baxerdale." Center for the Study of Organizational Innovation: University of Pennsylvania Discussion Paper No. 43, March 1979.

__________. "The Forward Market Effects of the Excuse Doctrines: Once More Unto the Breach." Center for the Study of Organizational Innovation: University of Pennsylvania Discussion Paper No. 59, November 1979.

Rogerson, William P. "Aggregate Expected Consumer Surplus as a Welfare Index with an Application to Price Stabilisation." Econometrica 48 (1980):423-436.

Schwartz, Alan. "The Case for Specific Performance." Yale Law Journal 89 (1979):271-306.

Shavell, Steven. "Breach of Contract, Damage Measures and Resource Allocation." Mimeo, 1978.

__________. "Risk Sharing and Incentives in the Principal and Agent Relationship." Bell Journal of Economics 10 (spring 1979): 55-73.

CHAPTER 4

REPUTATION AND PRODUCT QUALITY

I INTRODUCTION

A large fraction of economic activity is organized on the basis of contracts which are either legally unenforceable or at least unenforceable in practice. As the actions of parties to some joint economic activity become more complex or difficult for third parties to verify, two costs of contracting begin to grow larger. First, the enforcement and litigation costs begin to grow. Second, the transactions costs of drafting a contract which adequately describes the contemplated actions grow larger. The first cost tends to produce the situation that even if a contract is legally enforceable, it is in practice non-enforceable because the size of litigation and enforcement costs relative to the gains from performance make the option of legal enforcement unattractive. The second cost tends to produce the situation that actions are not adequately specified to be legally enforceable. In both situations, parties to a contract find themselves in the position where they need not fear legal sanctions for breaching a contractual obligation, yet they often do honor these obligations even when their apparent interests lie in breaching.

Economists have traditionally rationalized such behavior in the context of a world of self-interested actors by reference to "reputation." Economic actors operating on the basis of self-interest may well honor non binding contracts that appear not to be in their short run interest if future opportunities depend on adequate performance of current obligations. In this case contracts merely provide parties with a mutual basis for understanding what their obligations to one another are. A growing literature is examining the micro-foundations of such behavior from a repeated games standpoint (Green, 1980; Radner, 1979; Rubinstein, 1979; Townsend, 1979). Less attention, however, has been paid to the other perspective from which economists typically analyze a problem -- behavior of markets as a whole.

This paper models behavior of a market where product quality can vary and consumers cannot determine this quality prior to purchase. The "contract" being considered is the firm's promise to supply a high quality good. Akerlof (1970) was one of the first to consider such a market. He pointed out that if firms could not be differentiated by consumers and if consumers did not believe that a firm would continue to produce at its past quality level, then firms would all produce the lowest quality of goods. A market of "lemons" would result. The starting point of this paper is the assumption that a firm can build a reputation -- firms can be differentiated by consumers and consumers believe that the quality of a firm's performance in the future will be related to its performance in the past. This situation has been extensively analyzed (Klein and

Leffler 1979; Dybvig and Spatt 1980) under the assumption that consumers can judge quality with complete accuracy once they have received the good. Under this assumption, firms which misrepresent quality have a lifetime of one period. These papers are thus in a sense on the economics of "fly by night operators."

A very large class of goods does not satisfy this assumption of perfect consumer accuracy. In particular, consumers are often capable of performing only very partial and vague evaluations of the quality of professional services they receive from doctors, lawyers, banks, mechanics, opticians, etc. Furthermore, the quality of a service from a given professional may vary from time to time. This combination of observer error and actual quality variance makes it difficult for consumers to correctly evaluate the quality of service that a firm produces. In such a market a producer can contemplate staying in the market over the long run even if he misrepresents the quality of his product. The quality elasticity of demand is no longer infinite, but a finite positive number. The quality elasticity of demand is larger to the extent that reputation functions in the market. Reputation means two things. A higher quality firm is more likely to experience repeat sales from its current customers. A higher quality firm is also more likely to receive new customers from the pool of customers dissatisfied with their current firm. Both factors mean that a higher quality firm will in the long run have a larger number of customers than a low quality firm.

A key characteristic of most of the professions in the service sector is that they tend to severely limit information flow to the consumers they service. An insightful study of this phenomenon describes it as follows.

> Professions in the service sector exercise extensive control over . . . the type and amount of information publicly disseminated about these services. Professional codes of ethics usually prohibit advertising, limit brand name identification, strongly discourage public evaluation of other professionals' work, and place limitations wherever possible on other public indications of the characteristics, quality, or price of the services provided." (Benham and Benham, 1975.)

It is therefore important to know how such limitations on information flow might affect such markets. The analysis of this paper suggests that one effect of this limitation of information is to reduce the quality elasticity of demand. This in turn reduces the number of high quality firms and the number of consumers patronizing them. Three factors contribute to this result.

First, as people become better judges of the quality of services they receive, the quit rate of low quality firms relative to high quality firms becomes larger. That is, dissatisfied customers are more often justifiably dissatisfied. Second, customers searching for a new firm are more likely to choose a high quality firm as information in the market increases. Third, as information and consumer sophistication decrease, consumers become less able to supply one another with detailed descriptions of the quality of the firms they patronize. Word-of-mouth advertising becomes based more and more on simple positive or negative recommendations. In such a situation an externality is associated with the consumer's choice of a rule when to quit his current firm. His choice of rule affects other consumers through their reliance on his recommendation. The individually rational decision of how harsh a judge to be does not take this extra marginal benefit into account and consequently the consumer is a less harsh judge of his current firm than is socially optimal. This translates into a lower than optimal quality elasticity of demand.

The reward to producing high quality goods is more customers. The reward to producing low quality goods is lower costs and thus a higher profit per good. This latter factor becomes less important as market price rises above the average variable cost of producing high quality goods because the relative cost savings to producing low quality goods becomes smaller. In a zero profits entry equilibrium price rises above average variable cost as fixed costs grow. Therefore, ceteris paribus, industries which exhibit higher fixed costs should experience less problems with misrepresentation of quality. This suggests that barriers to entry established by professional associations which take the form of large fixed costs such as training requirements or license fees may well act to insure product quality, even in the absence of effective quality monitoring by the association. The magnitude of the incentive created is easily determined from cost data and could therefore be empirically investigated.

II THE DEPARTURE RATE AND ARRIVAL RATE

Firms are assumed to be able to produce either high quality or low quality goods. Variables associated with high quality firms will be subscripted with an "h"; those associated with low quality firms will be subscripted with an "ℓ". There are N identical consumers, each consuming one unit of the good per period. Consumers can only identify the quality of the good they receive after purchasing it, and even then only probabilistically. After purchasing a high quality good, consumers mistakenly identify it as low quality with probability α; after purchasing a low quality good, consumers mistakenly identify it as high quality with probability β. Assume that $\alpha + \beta < 1$. This means that the consumer is more likely to believe that a good is of high quality if it is of high quality, than if it is of low quality.[1]

Let γ_h be the random variable describing the observed quality of a good for a consumer patronizing a high quality firm. Let γ_ℓ be the similar random variable for a consumer patronizing a low quality firm. Let h denote observation of a high quality good and ℓ denote observation of a low quality good. Then γ_h and γ_ℓ can be written as follows.

$$\gamma_h = \begin{cases} h, & \text{with probability } 1 - \alpha \\ \ell, & \text{with probability } \alpha \end{cases} \tag{1}$$

$$\gamma_\ell = \begin{cases} h, & \text{with probability } \beta \\ \ell, & \text{with probability } 1 - \beta \end{cases} \tag{2}$$

Each consumer receives a new observation on the firm he is patronizing each period. This allows him to update his estimate of the probability that the firm he is patronizing is producing the high quality good. Let p_h^t and p_ℓ^t be the random variables describing the consumer's estimate of his firm's quality after patronizing it for t periods. At the end of the first period, the consumer's estimate depends only on that period's observation. Let the function g_1 describe the consumer's estimation process.

$$p_h^1 = g_1(\gamma_h) \tag{3}$$

$$p_\ell^1 = g_1(\gamma_\ell) \tag{4}$$

Assume that the consumer estimates a higher probability that the firm is producing high quality goods if he observes a high quality good. Formally, assume that

$$g_1(h) \geq g_1(\ell). \tag{5}$$

For periods after the first period, the consumer updates his existing estimate. Let the set of functions $\{g_t\}_{t=2}^{\infty}$ describe this behavior

$$p_h^t = g_t(p_h^{t-1}, \gamma_h) \quad t = 2,3,... \tag{6}$$

$$p_\ell^t = g_t(p_\ell^{t-1}, \gamma_\ell) \quad t = 2,3,... \tag{7}$$

As for g_1 assume that observations of high quality produce at least as large estimates. Formally, for any $p \in [0,1]$

$$g_t(p,h) \geq g_t(p,\ell). \tag{8}$$

Also, assume that a higher last period estimate results in at least as high an estimate this period. Formally, for $p,q \in [0,1]$ and $p > q$, then

$$g_t(p,h) \geq g_t(q,h) \tag{9}$$

and

$$g_t(p,\ell) \geq g_t(q,\ell). \tag{10}$$

The consumer chooses a time dependent reservation probability, $q(t)$, such that he quits his current firm and finds a new supplier if and only if his probability estimate at time t falls below $q(t)$.

Specific analysis of the nature of q(t) is not important for this paper. (See the previously cited work on repeated games for some related work.) When the consumer leaves his current firm, he searches among the firms, using any information he can find from consumers and other sources. This process will be modelled in section IV. The consumer then begins consumption and quality estimation with the new firm he selects. Let B_h^t and B_ℓ^t be the probabilities of the consumer spending at least t periods with, respectively, a high and low quality firm. We can write

$$B_h^t = \Pr\{\bigwedge_{i=1}^{t}(p_h^i \geq q(i))\} \tag{11}$$

$$B_\ell^t = \Pr\{\bigwedge_{i=1}^{t}(p_\ell^i \geq q(i))\}. \tag{12}$$

Now define the numbers B_h and B_ℓ.

$$B_h = 1 + \sum_{t=1}^{\infty} B_h^t \tag{13}$$

$$B_\ell = 1 + \sum_{t=1}^{\infty} B_\ell^t . \tag{14}$$

In the context of this rather general structure, we can now draw a conclusion that will be used in the remainder of the paper. It turns out that if a fixed number of new consumers, n, choose a given firm every period, then in the long run the firm's expected number of consumers is n/B_h or n/B_ℓ, depending on whether it produces high or low quality goods. That is, in the long run, on average $1/B_h$ or $1/B_\ell$ of the firm's customers leave it every period. Let d_h and d_ℓ denote these departure rates.

$$d_h = 1/B_h \tag{15}$$

$$d_\ell = 1/B_\ell \tag{16}$$

Furthermore d_h is at least as small as d_ℓ. That is, high quality firms experience a lower departure rate than low quality firms. Proposition 1 summarizes this result.

Proposition 1:

(1) If n new consumers arrive at a firm each period, in the long run the expected number of consumers is n/B_h or n/B_ℓ, depending upon whether the firm produces high or low quality goods. Therefore in the long run on average the fraction d_h and d_ℓ, respectively, of a high and low quality firm's customers leave it.

(2) $d_h \leq d_\ell$.

(3) d_h is non decreasing in $(1 - \alpha)$.
d_ℓ is non decreasing in $(1 - \beta)$.

Proof:

See Appendix. □

Firms of the same quality have already been modelled as possessing the same departure rate. It is natural to assume that firms of the same quality also have the same number of new customers arriving on average. The number of arrivals per high quality firm should not necessarily equal the number of arrivals per low quality firm, however. Consumers can obtain some information about potential firms to patronize from other consumers or other sources of information such as government or consumer associations. Let A_h and A_ℓ be the average number of consumers arriving at high and low quality firms, respectively. Let G_h and G_ℓ be the number of high and low quality firms. Then let a denote the ratio of arrivals, where a is in $[1,\infty]$.

$$a = \frac{A_h/G_h}{A_\ell/G_\ell} \qquad (17)$$

If there were absolutely no information available, then a would be 1. Consumers would be equally likely to choose any firm. As information about firm quality becomes better, consumers become better able to differentiate between high and low quality firms, and a becomes larger. For the discussion in section VI, a will be assumed to be created by a particular process. For the purposes of the other sections it can be any constant.[2]

III EQUILIBRIUM FIRM SIZE

Although many of the equilibrium quantities such as firm size are expected values, they will be spoken of as certain for ease of exposition. Firms reach a constant size when their number of arrivals equals their number of departures. This is defined as their equilibrium size. The conclusion of this section is that the higher number of arrivals and lower departure rate for high quality firms results in the equilibrium size of high quality firms being larger than the equilibrium size of low quality firms.

Let N_h and N_ℓ be the number of consumers in high and low quality firms. Let D_h and D_ℓ be the number of consumers departing from high and low quality firms. Then by definition

$$D_h = d_h N_h. \tag{18}$$

$$D_\ell = d_\ell N_\ell. \tag{19}$$

Finally, by (17)

$$\frac{A_h}{A_\ell} = \frac{G_h}{G_\ell} a. \tag{20}$$

The above three equations describe the structure of consumer flows in the system. Consumers leave high and low quality firms at the rates of d_h and d_ℓ. The ratio of arrivals at high and low quality firms is described by (20). In equilibrium, since firm size is constant, arrivals equal departures. That is, our equilibrium conditions are

$$A_h = D_h \tag{21}$$

and

$$A_\ell = D_\ell . \tag{22}$$

Substitute (18) and (19) into (20) by using the equilibrium conditions to yield

$$\frac{N_h/G_h}{N_\ell/G_\ell} = a\frac{d_\ell}{d_h}. \tag{23}$$

Let x_h and x_ℓ denote the size of high and low quality firms and γ denote ad_ℓ/d_h. Then (23) is

$$\frac{x_h}{x_\ell} = \gamma. \tag{24}$$

The relative size of high and low quality firms depends on the parameter γ. How can γ be interpreted? It becomes larger as the relative number of arrivals begins to favor high quality firms (a goes up) or the relative departure rate favors high quality firms (d_h goes down relative to d_ℓ.) Therefore γ is a composite measure of the extent to which arrivals and departures reward high quality firms relative to low quality firms. The parameter γ is always in $[1,\infty]$. When γ equals 1 consumer arrivals and departures do not discriminate between high and low quality firms. As γ grows larger, high quality firms begin to experience more arrivals and fewer departures than low quality firms. This results in a larger size for high quality firms relative to low quality firms.

One final assumption is that all firms face a capacity constraint of k. Therefore in equilibrium

$$x_h = k. \tag{25}$$

The single market price adjusts so that the high quality firms produce at capacity. Low quality firms masquerading as high quality firms necessarily produce at less than capacity given that they sell at the same price. An equilibrium in which all low quality firms charged a lower price and also produced at capacity would no longer involve misrepresentation of product quality, since consumers could use price to differentiate between product quality prior to purchase. Some firms may well choose to produce a lower quality product at a lower price. However, this amounts to the firm exiting the market for high quality goods and entering a different market. In this paper I study the single market for high quality goods and incentives for misrepresentation within it. The question of how markets for various qualities relate to one another and how relative prices equilibrate is a separate problem. Understanding the behavior within each market is the first step towards dealing with this larger problem.

IV FIRM ENTRY AND PRODUCT QUALITY CHOICE

At the market price, p, every firm must decide whether to actually produce a high quality product, to produce a low quality product and misrepresent its quality, or to not produce at all. Assume that there is a total of I firms, indexed by i. Let c_h^i and c_ℓ^i be the constant marginal cost for firm i of producing, respectively, high and low quality units. Let $F + F_i$ be the fixed cost of production for firm i.[3]

When firms compare the relative profitability of these three courses of action, they assume that any action on their part will leave average firm size for high and low quality firms and other firms' behavior unaffected. It is also assumed that firms are concerned with maximizing their long run average profits. In other words, firms do not discount future profits.[4] This assumption allows analysis of the long run incentives to misrepresent product quality as opposed to the incentives for "fly by night" operators. This latter problem has been analyzed by others (Klein and Leffler, 1979).

The long run average profits to be accrued from producing the high quality good, Π_h^i, are

$$\Pi_h^i = px_h - c_h^i x_h - F - F_i \tag{26}$$

and from producing the low quality good are

$$\Pi_\ell^i = px_\ell - c_\ell^i x_\ell - F - F_i. \tag{27}$$

The firm chooses the course of action associated with the largest of the two numbers Π_h^i and Π_ℓ^i. By rearranging (26) and (27),

$$\Pi_h^i \geq \Pi_\ell^i \iff \frac{p - c_h^i}{p - c_\ell^i} \geq \frac{x_\ell}{x_h} . \tag{28}$$

Rearrange to yield

$$\Pi_h^i \geq \Pi_\ell^i \iff p \geq \frac{x_h c_h^i - x_\ell c_\ell^i}{x_h - x_\ell} . \tag{29}$$

Let p_i^* be defined by

$$p_i^* = \frac{x_h c_h^i - x_\ell c_\ell^i}{x_h - x_\ell} . \tag{30}$$

In accord with Klein and Leffler (1979), call p_i^* the quality guaranteeing price for firm i. At prices above p_i^* firm i produces high quality goods, while at prices below p_i^* it produces low quality goods. As a convention, assume the firm produces high quality goods when the market price is p_i^* as well. It is easy to see that p_i^* is greater than c_h^i from (29) so long as $x_h > x_\ell$.

The intuition behind this result is clear. Since fixed costs are the same for either mode of production, only variable costs are relevant in a comparison. If price equalled the variable cost of producing high quality, then there would be no advantage to generating more business by producing high quality goods. It would be more profitable to sell fewer goods, but make a profit on each. As price rises above c_h^i, the advantage to generating more business by producing high quality goods becomes more substantial until finally at some point it becomes large enough that producing high quality goods becomes the most profitable course of action.

Note that there is no necessary relation between how high cost a firm is and whether it is more inclined towards producing high or low quality goods without further assumptions. Suppose that we can order the firms so that c_h^i and c_ℓ^i both increase in i. Even in this case p_i^* may not increase in i. However, if we additionally assume that c_h^i increases in i more quickly than c_ℓ^i, then p_i^* increases in i. In particular, this includes the case where c_h^i/c_ℓ^i is a constant. Therefore under a fairly plausible assumption, higher cost firms are more likely than lower cost firms to produce lower quality goods.

Two points concerning the firms quality choice in long run equilibrium should be noted. First, in equilibrium $x_h/x_\ell = \gamma$, so as γ becomes larger, the quality guaranteeing price becomes smaller. Reputation affects firm size through two processes in this model. A firm producing high quality goods experiences a lower departure rate and a higher arrival rate. The number γ is a measure of the combined strength of these two factors. As γ becomes larger, reputation matters more and a firm is more likely to produce high quality products.

Second, a "perfectly efficient" leave rate or arrival rate would solve the problem by itself. That is, if d_h/d_ℓ was 0 or a was ∞, then p^* would be 0 and all prices above c_h^i would induce production of high quality goods. Therefore the fact that consumers cannot determine the quality of a good prior to purchase does not in and of itself create a market with low quality goods if firms have sufficiently low discount rates for future business. Difficulty in determining the quality of a good even after purchase adds an entirely new reason to expect production of low quality goods. A firm may have an incentive to produce low quality goods even in the presence of no discounting when such post-purchase observability problems exist.

Firm i therefore makes its production decision as follows.[5]

x_h, x_ℓ and p are such that	Firm i's decision
$\Pi_h^i \geq \Pi_\ell^i \wedge \Pi_h^i \geq 0$	Produce high quality
$\Pi_\ell^i > \Pi_h^i \wedge \Pi_\ell^i \geq 0$	Produce low quality
$\Pi_h^i < 0 \wedge \Pi_\ell^i < 0$	Do not produce

Therefore the number of high and low quality firms is determined as follows. Let #S denote the number of elements in the set S.

$$G_h = \#\{i : \Pi_h^i \geq \Pi_\ell^i \wedge \Pi_h^i \geq 0\} \tag{31}$$

$$G_\ell = \#\{i : \Pi_\ell^i > \Pi_h^i \wedge \Pi_\ell^i \geq 0\} \tag{32}$$

$$G = G_h + G_\ell \tag{33}$$

V MARKET EQUILIBRIUM

Equations (23), (25), (31), and (32) together with the fact that there are N consumers, determine market equilibrium. For convenience, I rewrite them here.

$$\frac{N_h/G_h}{N_\ell/G_\ell} = \gamma \tag{34}$$

$$N_h/G_h = k \tag{35}$$

$$G_h = \#\{i : \frac{p - c_h^i}{p - c_\ell^i} \geq \frac{N_h/G_h}{N_\ell/G_\ell} \wedge (p - c_h^i)\frac{N_h}{G_h} \geq F + F_i\} \tag{36}$$

$$G_\ell = \#\{i : \frac{p - c_h^i}{p - c_\ell^i} < \frac{N_h/G_h}{N_\ell/G_\ell} \wedge (p - c_\ell^i)\frac{N_\ell}{G_\ell} \geq F + F_i\} \tag{37}$$

$$N_h + N_\ell = N \tag{38}$$

These five equations determine the five variables N_h, N_ℓ, G_h, G_ℓ and p. I assume that a solution exists to the above.

To derive comparative statics it is convenient to rewrite these five equations as follows:

$$N_h - kG_h = 0 \tag{39}$$

$$N_\ell - (k/\gamma)G_\ell = 0 \tag{40}$$

$$N_h + N_\ell - N = 0 \tag{41}$$

$$G_h - f_h(p,F,k,\gamma) = 0 \tag{42}$$

$$G_h - f_\ell(p,F,k,\gamma) = 0 \tag{43}$$

where f_h and f_ℓ are defined by

$$f_h(p,F,k,\gamma) = \#\{i : \frac{p - c_h^i}{p - c_\ell^i} \geq \frac{1}{\gamma} \wedge (p - c_h^i)k \geq F + F_i\} \tag{44}$$

and $$f_\ell(p,F,k,\gamma) = \#\{i : \frac{p - c_h^i}{p - c_\ell^i} < \frac{1}{\gamma} \wedge (p - c_\ell^i)\frac{k}{\gamma} \geq F + F_i\} \tag{45}$$

Since f_h and f_ℓ only assume integer values they are not differentiable. Therefore it is not strictly correct to apply the algorithm of total differentiation to obtain comparative statics results. However, we can perform exactly the same algorithm for small discrete changes of the parameters. To avoid the extra notational complexity of this procedure, I will formally use differential notation and apply the total differentiation algorithm. However, strictly speaking some of the infinite differences should actually be interpreted as small discrete changes.

Some properties of f_h and f_ℓ are useful in deriving comparative statics results. For reference, they are gathered together in Proposition 2.

Proposition 2:

(1) f_h is non decreasing in $\{p,\gamma,k\}$

(2) f_h is non increasing in $\{F\}$

(3) f_ℓ is non decreasing in $\{k\}$

(4) f_h is non increasing in $\{F,\gamma\}$

(5) $(f_h + f_\ell)$ is non decreasing in $\{k,p\}$

(6) $(f_h + f_\ell)$ is non increasing in $\{\gamma,F\}$

Proof:

Obvious. □

The comparative statics results are now stated in Proposition 3.

Proposition 3:

(1) Effects of γ:

(i) N_h and G_h are non-decreasing in γ.

(ii) N_ℓ is non-increasing in γ.

(iii) The effect of γ on p and G_ℓ is indeterminate.

(2) Effects of F:

Suppose that $f_{\ell p} \underset{(\geq)}{\leq} \frac{f_{\ell F} f_{hp}}{f_{hF}}$.

Then

(i) N_h and G_h are non-decreasing (non increasing) in F.

(ii) N_ℓ and G_ℓ are non-increasing (non decreasing) in F.

As well it is always true that

(iii) p is non-decreasing in F.

(3) Effects of k:

(i) p is non-increasing in k.

(ii) The effects of k on N_h, G_h, N_ℓ, and G_ℓ are indeterminate.

(4) Effects of N:

(i) N_h and G_h are non-decreasing in N

(ii) The effect of N on N_ℓ and G_ℓ is indeterminate.

(iii) p is non-decreasing in N.

Proof:

See Appendix. □

Corollary 3-a:

(i) N_h is non-increasing in α and β.

N_ℓ is non-decreasing in α and β.

(ii) N_h is non-decreasing in α.

N_ℓ is non-increasing in α.

Proof:

(i) This follows from Proposition 3 - (1) and Proposition 1 - (3) and the definition of γ.

(ii) This follows from Proposition 3 - (1) and the definition of γ.

□

These comparative statics can be best understood by first recasting (34) - (38) in a more traditional "supply equals demand" framework. The demand is N and is fixed. Supply is the number of high quality firms, G_h, times the size of high quality firms, k, plus the number of low quality firms, G_ℓ, times the size of low quality firms, k/γ. Long run equilibrium supply is thus $kG_h + (k/\gamma)G_\ell$. Market equilibrium occurs when

$$N = kf_h(p,F,k,\gamma) + \frac{k}{\gamma}f_\ell(p,F,k,\gamma). \tag{46}$$

It is equation (46) that determines market price. Then G_h, G_ℓ, N_h and N_ℓ are determined by

$$N_h = kf_h(p,F,k,\gamma) \tag{47}$$

$$N_\ell = \frac{k}{\gamma}f_\ell(p,F,k,\gamma) \tag{48}$$

$$G_h = f_h(p,F,k,\gamma) \tag{49}$$

$$G_\ell = f_\ell(p,F,k,\gamma). \tag{50}$$

Therefore, the effect of any parameter on price is determined solely by (46). The effect of any parameter on the other variables, N_h, N_ℓ, G_h, or G_ℓ, is determined by two factors -- first, its effect on price and price's subsequent effect on the other variable and, second, the direct effect of the parameter on the variable -- through the appropriate equation of (47) - (50). For example, we can write, based on (47)

$$\frac{\partial N_h}{\partial F} = k\frac{\partial f_h}{\partial p}\frac{\partial p}{\partial F} + \frac{k\partial f_h}{\partial F}. \tag{51}$$

The effect of F on equilibrium price is $\partial p/\partial F$ as determined by (46). Then $k\ (\partial f_h/\partial p)$ measures the effect of price on N_h. Finally, $k(\partial f_h/\partial F)$ measures the direct effect of F on N_h.

Effects of γ:

The effect of an increase of γ on supply is ambiguous. When γ increases, f_h increases and f_ℓ decreases. The total number of firms, $f_h + f_\ell$, decreases, but because more of them are the larger high quality firms, it is not clear whether $kf_h + (k/\gamma)f_\ell$ decreases or increases. Therefore the effect of γ on equilibrium price is also ambiguous. If γ increases supply, it decreases price; if γ decreases supply it increases price.[6]

Now consider the effect of γ on the other four variables. We can divide the effect, conceptually, into two parts: the direct effect of γ on the variable and the indirect effect through changes in the equilibrium price. First consider the direct effect -- assume that price is constant. Now consider the number of firms, G_h

and G_ℓ. The fact that γ goes up means that the relative size of high quality firms increases. Therefore every firm in the market finds it relatively more profitable to produce high quality goods and some firms which formerly produced low quality goods instead now produce high quality goods. Also, some firms which formerly found it profitable to be in the market as low quality producers now may find it unprofitable to be in the market at all. That is, when γ rises, the long run effect is to lower the size of low quality firms in absolute terms as well as relative to the size of high quality firms. Therefore some low quality firms may either begin producing high quality goods or leave the market entirely.

Now consider the direct effect of γ on N_h and N_ℓ (i.e., still assume that p is constant.) Even if G_h and G_ℓ remained constant, an increase in γ means that both relative arrival and departure rates favor high quality firms more and therefore the relative size of high quality firms increases. This means that N_h increases and N_ℓ decreases. The fact that G_h also goes up and G_ℓ goes down amplifies this change.

But p changes as well, to satisfy the equilibrium condition that supply equal demand. It can be shown that the indirect effects do not overwhelm the direct effects for N_h, G_h and N_ℓ; the effect on G_ℓ may be reversed. However, the equilibrium values of N_h and N_ℓ are probably of most interest since these are the number of high and low quality goods sold. Proposition 3 therefore states that when γ increases, the number of high quality units being sold increases and the number of low quality units being sold decreases. Recall that γ is a measure of the extent to which reputation operates in the market.

Reputation includes two phenomena: the tendency for a higher quality firm to receive more repeat purchases and the tendency for a higher quality firm to receive more new arrivals through word-of-mouth advertising. Both operate to affect relative firm size and γ measures their joint effect.

From Corollary 3-a, improving the consumer's ability to judge the quality of products and providing more information about the quality of all firms will, through their effect on γ, increase N_h and decrease N_ℓ. To the extent that advertising improves consumers' ability to judge and compare quality and to select new high quality firms, it should therefore result in a larger fraction of high quality products.

Effects of F:

The effect of an increase in the fixed costs of all firms on price is determinate. An increase in fixed costs does not change the relative profitability of high and low quality production, but it does make production in general less profitable. Therefore, the long run supply decreases and equilibrium price increases.

Now consider the effect of increasing F on the other four variables. The direct effect is to decrease all four, since firms of both types leave the industry. However, price then rises to bring supply back up to N and this also influences the four variables. The net effect on the number of each type of firm is therefore the sum of the effect from lowering F and then raising p so that total supply remains unchanged.

As price rises, all firms find production of high quality goods becoming more profitable relative to low quality goods. Therefore some existing firms in the market switch from production of low quality to high quality goods. However, because price rises there is an entry of firms producing both high and low quality goods. If this second factor is large enough, it is conceivable that as price goes up, the number of low quality firms as well as the number of high quality firms increases.

Suppose, for a moment, that this is not the case -- that the number of low quality firms decreases when price rises. Then the net effect of F is easy to sign. Raising F directly causes both G_h and G_ℓ to drop. Then the compensating price rise causes further drops in G_ℓ; however production returns to its original level and firm size is unchanged. Therefore G_h must have risen above its original level. In this case, an increase in fixed costs causes an increase in the number of high quality firms and a decrease in the number of low quality firms. Since F does not affect firm size, an increase in the number of consumers patronizing high quality firms and a decrease in the number of consumers patronizing low quality firms also results.

Even if a price rise causes an increase in the number of low quality firms, the above result will still hold as long as an increase in price does not cause G_ℓ to rise "too much." The initial increase in F causes both G_ℓ and G_h to drop. The equilibrium adjustment of price upward then causes G_ℓ and G_h to both rise. So long as the relative rise in G_ℓ is exceeded by the relative rise in G_h, the result still holds. That is, when $f_{\ell p}$ is positive we must have

$$\left|\frac{f_{\ell p}}{f_{\ell F}}\right| \leq \left|\frac{f_{hp}}{f_{hF}}\right| \tag{52}$$

for an increase in F to cause an increase in the equilibrium values of G_h and N_h and a decrease in the equilibrium values of G_ℓ and N_ℓ.

This condition is fairly plausible. For example suppose that entry and exit into the industry occur by low quality firms. That is, high quality firms are also the low cost, high profit firms. Section IV discussed conditions under which this might occur. Then $f_{hF} = 0$ but f_{hp} is positive due to "switchovers" by existing firms and the RHS of (52) is ∞. However, the reverse results may hold if an industry is such that low profit firms tend to be high quality producers and price rises cause very few switchovers from low to high quality but do cause entry of low quality firms.

This analysis suggests that industry organizations may well promote product quality even if they engage in little or ineffective quality monitoring. To the extent that they create large fixed costs for members by requiring training or license fees, production of high quality goods becomes more attractive to members. Furthermore, to the extent that the organization is able to act like a monopolist, it will restrict entry and raise prices, which once again creates an indirect incentive to produce high quality goods. This is not to suggest that anti-competitive trade associations are necessarily social benefactors. It may well be that the social costs of such anti-competitive actions outweigh the benefits of improved product quality. This is an empirical issue. The magnitude of the incentive for firms to produce high quality goods created by license fees or training costs for various industries is readily determinable from cost data. It would be interesting to determine if industry association requirements actually tend to substantially alter members' incentives or rather merely create opportunity for monopoly profits.

Effects of k:

Most effects of changing capacity are indeterminate. When capacity is increased, long run supply shifts out so equilibrium price drops.

Effects of N:

When the number of consumers increases, equilibrium price of course rises. The effect of N on the other four variables -- N_h, N_ℓ, G_h, G_ℓ -- is totally through the effect of increasing price. Therefore the number of high quality firms and the number of consumers patronizing high quality firms increases. The effect of price on the number of low quality firms is indeterminate for reasons previously discussed.

VI CHOICE OF A RESERVATION PROBABILITY

The consumer must decide how harsh a judge to be of his current firm when contemplating switching firms. The consumer would like to select some rule which tends to maximize his long run probability of patronizing a high quality firm. This goal is complicated by another factor -- search is costly. Therefore, the consumer might choose a rule involving less search than that which maximizes his probability of patronizing a high quality firm in order to save on search costs.

A more specific assumption about the process of consumer search for new firms is introduced in this section. Assume that consumers receive incomplete information from their friends and associates about the firms they patronize. In particular, consumers receive a judgment that the firm is either high quality or low quality as opposed to a detailed description of the friend's experience with a particular firm. A consumer can recommend either for or against his current firm when a fellow consumer makes an inquiry. He recommends for it if he intends to patronize it again next period. He recommends against it if he intends to switch himself.

This assumption probably becomes more reasonable when quality has many dimensions and is not easily summarized by any statistic. For example, the quality of a doctor involves a range of poorly measurable factors such as his diagnostic judgment, surgical competence, bedside manner, etc. The quality of many professional services has a number of dimensions. It may well be that in such a case the most often communicated piece of information is simply whether the firm is "good" or "bad."

One other necessary condition for this assumption to be reasonable is that consumers be relatively homogeneous, or at least that consumers be able to easily classify other consumers as to whether their tastes are similar. In the extreme case, we could imagine a case where quality consists of a number of different factors valued differently by different consumers. Firms are "stochastically identical" -- the fact that one consumer prefers firm 1 to firm 2 gives us no information about any other consumer's preferences. In this case, a positive or negative quality recommendation provides no information at all.[8]

The interest in analyzing this class of markets, as opposed to those where all available information is perfectly communicated, lies in the fact that an externality exists in the consumer's choice of how harsh a judge to be. When the consumer chooses a reservation probability he has an effect on those consumers who will use his judgment as a basis upon which to select a new firm. However, he does not take this into account when choosing a reservation probability. That is, there is an extra marginal benefit to choosing a higher reservation probability that the consumer ignores -- the benefit that other consumers receive from more reliable judgments. As a consequence the individually rational choice of a reservation probability is lower than the socially rational choice.

To demonstrate the externality formally, assume that each consumer uses the following sort of rule for deciding when to leave a firm he patronizes. He chooses a reservation probability, q.[9] The consumer continues to patronize the same firm until his estimate of the probability that the firm is high quality drops below q. Then he seeks out a new firm, possibly asking friends for recommendations. Recommendations are merely positive or negative. No actual probability estimate is communicated.

Consistent with the long run equilibrium orientation of this paper, we define an optimal reservation probability as the one which maximizes the consumer's long run average expected value. To define the socially and individually rational reservation probability some notation must be introduced. We consider a particular consumer and label all variables associated with him with a "*". Variables for the other consumers are identical and are denoted by the variable without a "*". We define a function $W(q^*,q)$ which yields the long run expected value to the particular consumer if he chooses q^* and everyone else chooses q. The particular consumer chooses an individually rational value of q^*. Let $Q(q)$ denote the particular consumer's choice of q^* given that everyone else chooses q. Then $Q(q)$ satisfies

$$W(Q(q),q) \geq \max_{\theta \in [0,1]} W(\theta,q). \tag{53}$$

However, the particular consumer was chosen arbitrarily and consumers are alike. Therefore, the value of q that consumers choose satisfies

$$q = Q(q). \tag{54}$$

The optimal level of q is the one such that the return to each consumer is maximized. That is, the optimal level of q satisfies

$$W(q,q) \geq \max_{\theta \in [0,1]} W(\theta,\theta). \tag{55}$$

The purpose of this section is to show that values of q which satisfy (54) are different, and in general smaller than, values of q which satisfy (55).

To do this, W must be defined. The departure rates for any consumer are a function of his choice of a reservation probability. Therefore we can write for some functions ξ_h and ξ_ℓ:

$$d_h = \xi_h(q)$$

$$d_\ell = \xi_\ell(q)$$

$$d^*_h = \xi_h(q^*)$$

$$d^*_\ell = \xi_\ell(q^*) \quad (56)$$

A crucial realization for the purposes of this section is that the relative arrival rate for the particular consumer, a^*, depends on d_h and d_ℓ but <u>not</u> d^*_h or d^*_ℓ. The consumer's choice about which firm to patronize next period, given that he has left his old firm, depends on what other consumers tell him about whether they intend to patronize their firms next period. This latter fact depends on d_h and d_ℓ. Therefore so does a^*. By composition with (56), a^* is a function of q but <u>not</u> q^*. Therefore we can write for some function, δ,

$$a^* = \delta(q). \quad (57)$$

An example may make this point clearer: Suppose that the consumer considers a firm by asking one of its patrons for a recommendation and accepts a firm as soon as he receives a positive recommendation. Suppose, as well, that the consumer has an equal chance of considering any firm. Then a^* is $1 - d_h/1 - d_\ell$.

Costs of search for the consumer rise along with q^* since a large reservation probability means the consumer will be departing from firms more often and thus searching more often. Let $S(q^*)$ be the consumer's search costs. Let u_h and u_ℓ be, respectively, the value of a high and low quality product to every consumer. The consumer now knows everything necessary to calculate an optimal reservation probability, except the long run probabilities of his patronizing a high quality firm or a low quality firm. Let h^* and ℓ^* denote these quantities.

Lemma 1:

$$h^* = \frac{d_\ell^* a^* G_h}{d_\ell^* a^* G_h + d_h^* G_\ell} \tag{58}$$

$$\ell^* = \frac{d_h^*(1 - G_h)}{d_\ell^* a^* G_h + d_h^* G_\ell} \tag{59}$$

Proof:

See Appendix. □

From Section V, G_h and G_ℓ are generally functions of γ and thus of q and q*. When G_h and G_ℓ vary with the consumers' choice of q other externalities than the one identified at the start of this section could be identified based on the fact that the consumer does not consider the benefit that accrues to others when he generates a higher G_h through search. However, this is a more standard argument. The externality originating because recommendations as opposed to objective quality information are communicated seems more novel and is thus the one I consider. To avoid confusing this argument with externalities involved with G_h and G_ℓ, I assume G_h and G_ℓ are constant.[10] Therefore h^* is a function of q and q* because d_h^* and d_ℓ^* depend on q* through (56) and because a^* depends on q through (57). Therefore we can write for some function, θ,

$$h^* = \theta(q^*,q). \tag{60}$$

Now W(q*,q) can be defined.

$$W(q^*,q) = \theta(q^*,q)u_h + (1 - \theta(q^*,q))u_\ell - S(q^*). \tag{61}$$

Assume that W(q*,q) is concave and differentiable in q* and that W(q,q) is concave and differentiable in q. We can now state the conclusion of this section.

Proposition 4:

Let x satisfy (54) and y satisfy (55). Then $x \neq y$ so long as $\theta_2(x,x) \neq 0$. Furthermore, if $\theta_2(q,q) > 0$ for every q, then $x < y$.

Proof:

See Appendix. □

When the individual chooses a reservation probability, he does not consider the benefit he confers on others through providing recommendations. That is, he ignores $\theta_2(q,q)$. Therefore so long as this benefit is non zero, he will in general make the wrong choice. It might be most reasonable to expect the relative arrival rate, a, to always grow with q. Then θ_2 is positive. This means that as consumers become harsher critics of firm performance they become relatively better at differentiating high quality firms from low quality firms. If this is the case, then the socially rational reservation probability is larger than the individually rational probability. When the individual chooses a reservation probability, he does not consider the benefit he confers on others through providing recommendations, and thus chooses too low a level. If θ_2 exhibits perverse behavior and becomes negative at some point, then the consumer may well be ignoring a negative externality and consequently choose q* too high. In either case, an externality exists which precludes achievement of a decentralized social optimum.

Of course this externality becomes smaller to the extent that consumers are able to communicate the actual estimated quality of their firm, as opposed to a positive or negative recommendation. This analysis therefore reinforces the policy recommendations of Wilde and Schwartz (1979), although for entirely different reasons.

> This suggests that the state should reduce the costs to consumers of comparing purchase alternatives. One way to achieve this is for the state to require more standardization of the way in which firms quote prices and terms because such standardization would reduce the cost of comparison shopping. It also seems wise to remove barriers to private, voluntary standardization. Thus courts should not regard the use by a seller of a standard form contract as a factor which militates against enforcement of the contract, for such judicial conduct raises the cost to firms of creating standardized forms. Also, legislatures should consider relaxing antitrust enforcement to permit more voluntary standardization of the ways in which prices and terms are quoted.

Also, to the extent that advertising generally increases the sophistication and awareness of all consumers, it may facilitate transfer of information among consumers.

VI CONCLUSION

When consumers have difficulty in accurately evaluating the quality of a good or service, firms may contemplate staying in the market over the long run even though they misrepresent quality. "Reputation" provides an incentive for firms to produce high quality items. Through differential quit and arrival rates, higher quality firms tend to have more customers. The model seems to apply particularly well to professional services. Professional associations tend to severely limit information flow to the consumers they service. Three different factors suggest that this lowers the quality elasticity of demand and thus results in higher fractions of low quality firms and consumers patronizing them. One of the three factors depends on the observation that when transaction costs severely limit word-of-mouth advertising an externality begins to be attached to the consumer's choice of whether or not to continue patronizing a firm. I believe this idea is new to the literature.

The formal model in this paper resembles in some respects that of Schmalensee (1978), in which he investigates the relationship between advertising and product quality. Product quality is fixed in his model and advertising is the choice variable of the firms. It is natural to ask if Schmalensee's results are robust to the case where firms may also choose product quality as in this paper's model. The other, more difficult, direction for research suggested by this paper concerns the relationship between price dispersion and product quality. This paper assumes that a single market price exists. Presumably another reward for establishing a reputation might be the ability to charge a higher price.

APPENDIX

Proposition 1:

(1) Consider a high quality firm's population of customers. There are n customers who just arrived. There are an average of $B_h^1 n$ customers who have been with the firm one period before. In general there are an average of $B_h^t n$ customers who have been with the firm t periods before. Therefore the firm's total number of customers is

$$\left(\sum_{t=1}^{\infty} B_h^t n\right) + n. \tag{A-1}$$

Therefore the total number of consumers is nB_h. If n customers arrive every period and the expected number of customers stays constant at nB_h then the expected number of customers that leave each period is also n. That is, on average d_h of the firm's total customers leave each period. The case for a low quality firm is obviously similar.

(2) It is clearly sufficient to show that $B_h^t \geq B_\ell^t$ for every t. Choose any t and fix it. Let Γ_h and Γ_ℓ be the random vector of t repetitions of, respectively, γ_h and γ_ℓ.

$$\Gamma_h = (\gamma_h, \gamma_h, \ldots, \gamma_h) \tag{A-2}$$

$$\Gamma_\ell = (\gamma_\ell, \gamma_\ell, \ldots, \gamma_\ell) \tag{A-3}$$

Let Δ^n be $\{h, \ell\}^n$. Therefore Δ^t is the range of Γ_h and Γ_ℓ. Let $\delta \in \Delta^t$. We will write $\delta = (\delta_1, \delta_2, \ldots, \delta_t)$. Let P_h and P_ℓ be the random vectors of the first t probability estimates.

$$P_h = (P_h^1, \ldots, P_h^t) \tag{A-5}$$

$$P_\ell = (P_\ell^1, \ldots, P_\ell^t) \tag{A-6}$$

Based on (3), (4), (6), and (7), recursively define the function $G\colon \Delta^t \to [0,1]^t$:

$$P_h = G(\Gamma_h) \tag{A-7}$$

$$P_\ell = G(\Gamma_\ell) \tag{A-8}$$

Because of the monotonicity assumption (5), (8), (9), and (10), it is clear that G is also monotone. That is if δ and $\delta^* \in \Delta$ and $\delta_i = h \Rightarrow \delta_i^* = h$ then

$$G(\delta^*) \geq G(\delta). \tag{A-9}$$

(For two vectors $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$ in R^n we say that $x \geq y$ if $x_i \geq y_i$ for every $i = 1, \ldots, n$.)

Now let q be any vector in $[0,1]^t$; q is the vector of reservation probabilities.

$$q = (q_1, q_2, \ldots, q_t) \tag{A-10}$$

We want to prove that

$$\Pr\{G(\Gamma_h) \geq q\} \geq \Pr\{G(\Gamma_\ell) \geq q\}. \tag{A-11}$$

To do this, construct a set of random vectors $\{\Gamma^i\}_{i=0}^t$. Each random vector is a t-tuple

$$\Gamma^i = (\Gamma_1^i, \Gamma_2^i, \ldots, \Gamma_t^i) \tag{A-12}$$

where

$$\Gamma_j^i = \begin{cases} \gamma_h, & j \leq i \\ \gamma_\ell, & j > i. \end{cases} \tag{A-13}$$

By definition, Γ^o is Γ_ℓ and Γ^t is Γ_h. Therefore, to prove (A-11) it is sufficient to show that for i=0,...,t-1

$$\Pr\{G(\Gamma^{i+1}) \geq q\} \geq \Pr\{G(\Gamma^i) \geq q\}. \tag{A-14}$$

For notational convenience it is easiest to show this fact for i = 0. However, it will be clear that the method of proof is perfectly general. Rewrite the random vectors $G(\Gamma^1)$ and $G(\Gamma^o)$ as a discrete combination of two random variables.

$$G(\Gamma^1) = \begin{cases} G(h,\gamma_\ell,\ldots,\gamma_\ell), & \text{with probability } (1 - \alpha) \\ G(\ell,\gamma_\ell,\ldots,\gamma_\ell), & \text{with probability } \alpha. \end{cases} \tag{A-15}$$

$$G(\Gamma^o) = \begin{cases} G(h,\gamma_\ell,\ldots,\gamma_\ell), & \text{with probability } \beta \\ G(\ell,\gamma_\ell,\ldots,\gamma_\ell), & \text{with probability } (1 - \beta) \end{cases} \tag{A-16}$$

Now define the following subsets of Δ^{t-1}.

$$\begin{aligned} H &= \{\delta \in \Delta^{t-1} : G(h,\delta) \geq q \wedge G(\ell,\delta) \not\geq q\} \\ L &= \{\delta \in \Delta^{t-1} : G(h,\delta) \geq q \wedge G(\ell,\delta) \geq q\} \\ N &= \{\delta \in \Delta^{t-1} : G(h,\delta) \not\geq q \wedge G(\ell,\delta) \not\geq q\} \end{aligned} \tag{A-17}$$

It is clear from (A-9) that H, L, and N are disjoint and their union is Δ^{t-1}. Finally, let Γ_ℓ^{t-1} denote the random vector of (t-1)-triple of γ_ℓ.

$$\Gamma_\ell^{t-1} = \underbrace{(\gamma_\ell,\gamma_\ell,\ldots,\gamma_\ell)}_{t\,-\,1\text{ times}} \tag{A-18}$$

Based on this notation, we can rewrite the LHS of (A-14) for i=0 as

$$(1 - \alpha)\Pr\{\Gamma_\ell^{t-1} \in H\} + \Pr\{\Gamma_\ell^{t-1} \in L\}. \tag{A-19}$$

We can rewrite the RHS of (A-14) for i = 0 as

$$\beta \Pr\{\Gamma_{\ell}^{t-1} \varepsilon H\} + \Pr\{\Gamma_{\ell}^{t-1} \varepsilon L\}. \tag{A-20}$$

Because $\beta \leq 1 - \alpha$, (A-20) is less than or equal to (A-19). □

Proposition 3:

Totally differentiate the system of equations (39) - (43) to yield

$$\begin{bmatrix} 1 & 0 & -k & 0 & 0 \\ 0 & 1 & 0 & -k/\gamma & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & -f_{hp} \\ 0 & 0 & 0 & 1 & -f_{\ell p} \end{bmatrix} \begin{bmatrix} dN_h \\ dN_{\ell} \\ dG_h \\ dG_{\ell} \\ dp \end{bmatrix} = \begin{bmatrix} 0 & G_h & 0 & 0 \\ 0 & G_{\ell}/\gamma & \frac{-kG_{\ell}}{\gamma^2} & 0 \\ 0 & 0 & 0 & 1 \\ f_{hF} & f_{hk} & f_{h\gamma} & 0 \\ f_{\ell F} & f_{\ell k} & f_{\ell \gamma} & 0 \end{bmatrix} \begin{bmatrix} dF \\ dk \\ d\gamma \\ dN \end{bmatrix} \tag{A-21}$$

$$\begin{bmatrix} dN_h \\ dN_{\ell} \\ dG_h \\ dG_{\ell} \\ dp \end{bmatrix} = \frac{1}{kf_{hp} + \frac{k}{\gamma}f_{\ell p}} \begin{bmatrix} \frac{k}{\gamma}f_{\ell p} & -kf_{hp} & kf_{hp} & \frac{k^2}{\gamma}f_{\ell p} & -\frac{k^2}{\gamma}f_{hp} \\ -\frac{k}{\gamma}f_{\ell p} & kf_{hp} & \frac{k}{\gamma}f_{\ell p} & -\frac{k^2}{\gamma}f_{\ell p} & \frac{k^2}{\gamma}f_{hp} \\ -f_{hp} & -f_{hp} & f_{hp} & \frac{k}{\gamma}f_{\ell p} & \frac{-k}{\gamma}f_{hp} \\ -f_{\ell p} & -f_{\ell p} & f_{\ell p} & -kf_{\ell p} & kf_{hp} \\ -1 & -1 & +1 & -k & -\frac{k}{\gamma} \end{bmatrix} \begin{bmatrix} 0 & G_h & 0 & 0 \\ 0 & G_{\ell}/\gamma & \frac{-kG_{\ell}}{\gamma^2} & 0 \\ 0 & 0 & 0 & 1 \\ f_{hF} & f_{hk} & f_{h\gamma} & 0 \\ f_{\ell F} & f_{\ell k} & f_{\ell \gamma} & 0 \end{bmatrix} \begin{bmatrix} dF \\ dk \\ d\gamma \\ dN \end{bmatrix} \tag{A-22}$$

Therefore we have

$$\begin{bmatrix} dN_h/d\gamma \\ dN_\ell/d\gamma \\ dG_h/d\gamma \\ dG_\ell/d\gamma \\ dp\ /d\gamma \end{bmatrix} = \frac{1}{kf_{hp} + \frac{k}{\gamma}f_{\ell p}} \begin{bmatrix} \frac{k^2}{\gamma}\{f_{h\gamma}f_{\ell p} - f_{\ell\gamma}f_{hp}\} + \frac{k^2 G_\ell f_{hp}}{\gamma^2} \\ \frac{k^2}{\gamma}\{f_{h\gamma}f_{\ell p} - f_{\ell\gamma}f_{hp}\} - \frac{k^2 f_{hp} G_\ell}{\gamma^2} \\ \frac{k}{\gamma}\{f_{h\gamma}f_{\ell p} - f_{\ell\gamma}f_{hp}\} + \frac{kf_{hp}G_\ell}{\gamma^2} \\ -k\{f_{h\gamma}f_{\ell p} - f_{\ell\gamma}f_{hp}\} + \frac{kf_{\ell p}G_\ell}{\gamma^2} \\ -k\{f_{h\gamma} + f_{\ell\gamma}/\gamma\} + \frac{kG_\ell}{\gamma^2} \end{bmatrix} \tag{A-23}$$

$$\begin{bmatrix} dN_h/dF \\ dN_\ell/dF \\ dG_h/dF \\ dG_\ell/dF \\ dp\ /dF \end{bmatrix} = \frac{1}{kf_{hp} + \frac{k}{\gamma}f_{\ell p}} \begin{bmatrix} \frac{k^2}{\gamma}\{f_{hF}f_{\ell p} - f_{\ell F}f_{hp}\} \\ -\frac{k^2}{\gamma}\{f_{hF}f_{\ell p} - f_{\ell F}f_{hp}\} \\ \frac{k}{\gamma}\{f_{hF}f_{\ell p} - f_{\ell F}f_{hp}\} \\ -k\{f_{hF}f_{\ell p} - f_{\ell F}f_{hp}\} \\ -k\{f_{hF} + f_{\ell F}/\gamma\} \end{bmatrix} \tag{A-24}$$

$$\begin{bmatrix} dN_h/dk \\ dN_\ell/dk \\ dG_h/dk \\ dG_\ell/dk \\ dp\ /dk \end{bmatrix} = \frac{1}{kf_{hp} + \frac{k}{\gamma}f_{\ell p}} \begin{bmatrix} f_{\ell p}\{\frac{k}{\gamma}G_h + \frac{k^2}{\gamma}f_{hk}\} - f_{hp}\{\frac{k}{\gamma}G_\ell + \frac{k^2}{\gamma}f_{\ell k}\} \\ -f_{\ell p}\{\frac{k}{\gamma}G_h + \frac{k^2}{\gamma}f_{hk}\} + f_{hp}\{\frac{k}{\gamma}G_\ell + \frac{k^2}{\gamma}f_{\ell k}\} \\ -f_{hp}\{G_h + \frac{G_\ell}{\gamma} + \frac{k}{\gamma}f_{\ell k}\} + f_{\ell p}\{\frac{k}{\gamma}f_{hk}\} \\ -f_{\ell p}\{G_h + \frac{G_\ell}{\gamma} + kf_{hk}\} + f_{hp}\{kf_{\ell k}\} \\ -G_h - \frac{G_\ell}{\gamma} - kf_{hk} - \frac{k}{\gamma}f_{\ell k} \end{bmatrix} \tag{A-25}$$

$$\begin{bmatrix} dN_h/dN \\ dN_\ell/dN \\ dG_h/dN \\ dG_\ell/dN \\ dp/dN \end{bmatrix} = \frac{1}{kf_{hp} + \frac{k}{\gamma}f_{\ell p}} \begin{bmatrix} kf_{hp} \\ \frac{k}{\gamma}f_{\ell p} \\ f_{hp} \\ f_{\ell p} \\ 1 \end{bmatrix} . \qquad \text{(A-26)}$$

Now a number of terms need to be signed to prove the Proposition. I will list each term and its sign as a claim and then prove it.

<u>Claim 1</u>: $kf_{hp} + \frac{k}{\gamma}f_{\ell p} \geq 0.$

<u>Proof</u>:

By Proposition 2 - (1), $f_{hp} \geq 0.$

By Proposition 2 - (5) $f_{hp} + f_{\ell p} \geq 0.$

As well, $\gamma \geq 1.$

Therefore

$$\begin{aligned} kf_{hp} + \frac{k}{\gamma}f_{\ell p} &\geq \frac{k}{\gamma}f_{hp} + \frac{k}{\gamma}f_{\ell p} \\ &= \frac{k}{\gamma}(f_{hp} + f_{\ell p}) \\ &\geq 0. \end{aligned} \qquad \text{(A-27)}$$

□

<u>Claim 2</u>: $f_{h\gamma}f_{\ell p} - f_{\ell\gamma}f_{hp} \geq 0.$

Proof:

By Proposition 2, $f_{hp} \geq 0$ and $f_{hp} + f_{\ell p} \geq 0$. Therefore either $|f_{hp}| \geq |f_{\ell p}|$ or $f_{\ell p} \geq 0$ is true.

By Proposition 2, $f_{h\gamma} \geq 0$, $f_{\ell\gamma} \leq p$, and $f_{h\gamma} + f_{\ell\gamma} \leq 0$. Therefore $|f_{\ell\gamma}| \geq |f_{h\gamma}|$.

Suppose that $|f_{hp}| \geq |f_{\ell p}|$. Then

$$\left.\begin{array}{l} |f_{hp}| \geq |f_{\ell p}| \\ |f_{\ell\gamma}| \geq |f_{h\gamma}| \end{array}\right\} \Rightarrow |f_{hp} f_{\ell\gamma}| \geq |f_{\ell p} f_{h\gamma}|. \qquad \text{(A-28)}$$

Since $-f_{hp} f_{\ell\gamma} \geq 0$, Claim 2 is true. Now suppose that $f_{\ell p} \geq 0$. Then Claim 2 is clearly true because both terms in the sum are non negative.

□

Claim 3: $f_{hF} f_{\ell p} - f_{\ell F} f_{hp} \geq 0$.

Proof:

The proof is similar to that of Claim 2. □

Proposition 3 now follows directly. □

Lemma 1:

Let h_t be the probability that the consumer patronizes a high quality firm at time t. The consumer's probability of leaving a high quality firm is d_h^*. His probability of leaving a low quality firm is d_ℓ^*. Given that he leaves his current firm, his probability of arriving at a high quality firm is $a^*G_h/(G_\ell + a^*G_h)$ and at a low quality firm is $G_\ell/(G_\ell + a^*G_h)$. Therefore, we can write

$$h_{t+1} = h_t \left[(1 - d_h^*) + \frac{d_h^* a^* G_h}{G_\ell + a^* G_h} \right] + (1 - h_t) \frac{d_\ell^* a^* G_h}{G_\ell + a^* G_h} \qquad \text{(A-29)}$$

This Markov process converges to h^*. The value for ℓ^* can be directly calculated in the same fashion or can simply be calculated by subtracting h^* from 1. □

Proposition 4:

The solution to (54) satisfies

$$\theta_1(x,x)(u_h - u_\ell) - S'(x) = 0 \qquad \text{(A-30)}$$

The solution to (55) satisfies

$$[\theta_1(y,y) + \theta_2(y,y)](u_h - u_\ell) - S'(y) = 0. \qquad \text{(A-31)}$$

These are different so long as $\theta_2(x,x) \neq 0$. By assumption the LHS of (A-30) slopes down. Therefore if $\theta_2(q,q)$ is positive for every q, the LHS of (55) intersects the axis at a point to the right of x.

□

FOOTNOTES

1. As explained in the introduction, this mathematics can also be interpreted to model a situation where a firm attempting to produce high (low) quality goods does not have complete control over the process and thus sometimes produces low (high) quality goods. Then the assumption that $\alpha + \beta < 1$ means that a firm attempting to produce high quality goods is more likely to produce high quality goods than is a firm attempting to produce low quality goods.

2. It is being assumed that firm size does not affect relative arrival rates. Nelson (1970) has correctly pointed out that larger firms might be better able to establish a reputation for a number of reasons. Such a modification should not substantially change the analysis.

3. Fixed costs are written as $F + F_i$ instead of merely as F_i so that the parameter F can be varied for comparative statics.

4. Schmalensee (1978) also uses this objective function to derive his comparative statics.

5. As a convention, the firm chooses to produce when the best production process yields zero profits and chooses to produce high quality when

production of either high or low quality yields identical non-negative profits.

6. This result depends on the supply curve sloping upward in price. This is proven in the Appendix in the proof to Proposition 3 under Claim 1.

7. Recall the discussion surrounding equation (29) concluded that the profits from high quality production relative to low quality production rise as price rises.

8. Satterthwaite (1979) considers such a model. Not surprisingly, he models information transfer as an exchange of facts about the specific aspects of a firm's quality. Satterthwaite also allows parties to exchange secondhand information and to possess memories of previous exchanges. He can manage such an elaborate model because it is not necessary for him to consider explicitly the nature of the information transferred, only the amount. This is a fruitful area for further research.

9. In previous sections q was allowed to depend on t. A reading of Rubinstein (1979) suggests that the optimal reservation probability will vary with t. For this section it is easier to analyze the optimal choice of the consumer if the domain is the real line. Instead of simply assuming the reservation probability is constant over time, we could view the consumer as already having performed

a partial optimization. Let q be associated with the best policy in the class of policies whose average is q.

10. The simplifying assumption is <u>not</u> that consumers assume G_h and G_ℓ to be constant when choosing a value of q. It is reasonable for consumers to exhibit behavior of this sort. This is similar to the behavior of firms assumed in Section III -- they assume their actions do not affect relative firm size. Both types of behavior derive from the fact that individual agents are a very small part of the market and the assumption of Nash behavior. The simplifying assumption being made in this section involves calculation of the optimal level of q. There are actually two separate issues involved in determining whether consumers choose an optimal level of q. One is whether the non-cooperative choice of q by a consumer adequately takes into account the benefit that he confers on all other consumers by changing G_h and G_ℓ. The other is whether non-cooperative choice of q by a consumer adequately takes into account the benefit that he confers on all other consumers through the recommendations he gives to them. I isolate the second issue by assuming that G_h and G_ℓ are constant when calculating an optimal level of q with which to compare the consumer's choice. It seems intuitive that the first factor would reinforce the second -- the optimal level of q is higher and involves more search than the chosen level. However, I have formally only analyzed the second factor.

REFERENCES

Ackerlof, George A. "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism." Quarterly Journal of Economics 84 (August 1970):488-500.

Benham, L. and Benham A. "Regulating the Professions: A Perspective on Information Control." Journal of Law and Economics 18 (1975):421-447.

Darby, Michael R. "Qualitative Information, Reputation, and Monopolistic Competition." Mimeo, March 4, 1976.

__________ and Karni, Ed. "Free Competition and the Optimal Amount of Fraud." Journal of Law and Economics 16 (April 1973):67-88.

Dybvig, Philip H. and Spatt, Chester S. "Does It Pay to Maintain a Reputation?" Mimeo.

Green, Edward. "Agency, Information and Cooperation over Time." Mimeo, February 1980.

Klein, Benjamin and Leffler, Keith B. "The Role of Price in Guaranteeing Quality." UCLA Discussion Paper No. 149.

Macauley, Stewart. "Non-Contractual Relations in Business: A Preliminary Study." American Sociological Review 28 (February 1963):55-69.

Nelson, Phillip. "Information and Consumer Behavior." Journal of Political Economy 78 (1970):311-329.

Radner, Roy. "Collusive Behavior in Noncooperative Epsilon-Equilibria of Oligopolies with Long but Finite Lives." Bell Labs Discussion Paper #144, August 1979.

Rothschild, Michael. "Models of Market Organization with Imperfect Information: A Survey." Journal of Political Economy 81 (December 1973):1238-1308.

Rubinstein, Ariel. "An Optimal Conviction Policy for Offenses that May Have Been Committed by Accident," Hebrew University Department of Economics Research Report No. 114, January 1979.

Salop, S. "Information and Monopolistic Competition." American Economic Review 66 (May 1976):240-245.

__________, and Stiglitz, J. "Bargains and Ripoffs: A Model of Monopolistically Competitive Price Dispersion." Review of Economic Studies 44 (1977):493-510.

Satterthwaite, Mark A. "Consumer Information, Equilibrium Industry Price, and the Number of Sellers." Bell Journal 10 (1979): 483-502.

Schmalensee, R. "A Model of Advertising and Product Quality." Journal of Political Economy 86 (1978):485-504.

Townsend, Robert M. "Contract Length and the Discipline of Enduring Relationships." Mimeo, July 1979.

Wilde, Louis L. and Schwartz, Alan. "Equilibrium Comparison Shopping." Review of Economic Studies 46 (July 1979):543-553.