

ESSAYS ON THE ECONOMICS OF CONTRACTS

by

LARS ANDREAS STOLE

B.A. (Political Science), University of Illinois-Champaign (1985)

B.A. (Economics), University of Illinois-Champaign (1985)

M.Sc., London School of Economics (1986)

SUBMITTED TO THE DEPARTMENT OF ECONOMICS
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE
DEGREE OF

DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 1991

©Lars Andreas Stole, 1991. All rights reserved.

The author hereby grants to MIT permission to reproduce and to
distribute copies of this thesis document in whole or in part.

Author

Department of Economics

May 7, 1991

Certified by

Jean Tirole

Professor of Economics

Thesis Supervisor

Accepted by

Richard Eckaus

Chairman, Departmental Committee on Graduate Students

MASSACHUSETTS INSTITUTE
OF TECHNOLOGY

Department of Economics

JUN 13 1991

ARCHIVES

LIBRARIES

Essays on the Economics of Contracts

by

Lars Andreas Stole

Submitted to the Department of Economics
on May 7, 1991, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

In this thesis, I analyze the economics of three different contracting problems. The first chapter considers the contracting problem facing multiple principals, each of whom desires to contract with the same agent. If the agent has private information regarding his gains from the contracting activity and the contracting activities in the principal-agent relationships are substitutable (complementary), the principals will typically extract less (more) information rents in total and induce less (more) productive inefficiency in the contracting equilibrium than if there were a single principal contracting over the same activities. This analysis is subsequently applied to various environments, including joint ventures, exclusive-dealing relationships, and regulation between conflicting governmental agencies.

The second chapter considers the potential use of liquidated damage clauses under asymmetric information. Courts typically allow parties to stipulate the damages each will pay the other in event of breach, providing that such liquidated damage terms *do not greatly exceed* actual losses. This restriction acts as a ceiling, however, as courts generally enforce terms that are equal to or *below* actual losses. This anomaly can be explained when bargaining occurs under asymmetric information. Here, the liquidated damage clause serves a dual role, both promoting efficient breach and signaling a party's valuation of trade. I show that it is always optimal for parties to set damages at or below valuation, thereby providing a consistent theory for the courts' asymmetric treatment of contractual damages: When damages are significantly below actual losses, courts may plausibly maintain the presumption that the contract is the result of rational bargaining. But when damages exceed losses, they must consider the likelihood of other elements such as mistake and fraud.

The final chapter examines the problem of procurement and second sourcing. When the government procures an item, ideally it would like to pay no more than the minimum possible cost. But in practice the government does not know all of the technological possibilities, even if it could perfectly audit most incurred costs. As a consequence, suppliers who know they have superior technology can earn extra profits – commonly known as “information rents”. A standard way to reduce information rents is to use competitive sealed-bid tenders, which in effect award the contract to the best producer at the second-best bidder's break-even cost. One problem with the preceding method

is that only one source may be technically qualified to bid. If government wants more competition, it must “generate” it through licensing or some other form of technology transfer. Conventional cost analysis wisdom asserts that transfer is merited if and only if the second source’s direct costs plus the costs of transfer are less than the first source’s direct costs. Chapter 3 shows that technology transfer offers an additional potential gain: reduction of information rents. To provide appropriate incentives, technology should perhaps be transferred even when the second source is less efficient than the first. Additionally, when developer moral hazard exists with respect to investments in cost-reducing technology, the optimal auction will make the developer’s success in the auction more sensitive to the developing firm’s announced costs.

Thesis Supervisor: Jean Tirole

Title: Professor of Economics

To my parents.

Acknowledgments

This dissertation is the culmination of over twenty years of learning, which in my case was possible only through the support and encouragement of many people.

First and foremost, I thank my father and mother for their love and support (financial and moral) through all of these years. I could not have attained what I have without them. My family has been very supportive and enthusiastic about my studies. My sisters, Joan and Marie, and my grandmother, have always been there for me.

I have also had the blessing of wonderful friends throughout my college and graduate studies: Tom Barnett, Kimi King, Lindsey Klecan, Laura McDermith, Phil and Sherri Messersmith, Miranda Oshige, Kevin Rakers, and Liz Read. Looking back on the past 9 years since I graduated from high school, I feel extremely fortunate having known these people; there were great times because of them. And to the extent that everyone of them at one time or another listened to my lamentations about economics or my doubts about being an economist, I owe to each of them a special debt.

Probably every scholar can point to some great teacher who initially inspired their studies. In my case, it was Franklin Shupp who inspired me while I was a sophomore at the University of Illinois. He showed me the poetry behind economic thought. Walter McMahon and Werner Baer were also very influential in my choice of economics as a career while I was in Champaign. Since my graduation from the University of Illinois, I have had other professors who have had a strong effect on my studies: Charlie Bean and John Moore at the London School of Economics, and David Pearce at Yale University.

In my first two years at MIT, I was fortunate to have four great teachers: Bob Gibbons, Drew Fudenberg, Oliver Hart, and Jean Tirole. I first learned contract theory from Bob Gibbons, who inspired me to work on rationalizability in signaling games for a term paper which later became joint work in the *Journal of Economic Theory* with Joel Sobel and Iñigo Zapater. Drew Fudenberg arrived at MIT in the fall of 1988, and I had the good fortune of taking several classes from him as well as having him as an advisor. Without Drew, this thesis would certainly have lacked in mathematical rigor and been less focused. I am also thankful to have had classes from Oliver Hart on the theory of

the firm and, my main field of interest, contract theory; Oliver also served as an advisor for my thesis and has been very helpful.

To Jean Tirole, I will always be indebted. From the start of my studies at MIT, Jean took a special interest in me and constantly encouraged me to tackle new interesting and difficult problems. Jean's generosity and insight have had a profound impact on everything I have done at MIT and I count myself as very fortunate to have been one of his students. I can only hope that in the future when I am working at the University of Chicago that I continue to do as good work without his supervision.

Another great attribute of MIT is the students (both classmates and visitors). I have learned almost as much from them as from formal classes. In particular, I am thankful for numerous conversations with Jim Dana, Tore Ellingsen, Keith Head, David Martimort, Kathy Spier, and Jeff Zwiebel, and I look forward to working with many of them in the future.

I have also had the good fortune of discussions with several people outside of MIT during my graduate studies. Specifically, I would like to thank Joel Demski, Kent Osband, David Sappington, and Jeremy Stein for helpful discussions, and Joel Sobel, Iñigo Zapater, and Lucian Bebchuk as valuable co-authors. Canice Prendergast, a classmate from the LSE and Yale (and now a colleague at the University of Chicago), has also proved to be a great resource.

Education is not free, and my education was no exception. Without my parents' substantial financial contributions, I would not have made it as far as I have. I would also like to thank the National Science Foundation, the Olin Foundation (both through Harvard's Program for Law and Economics and MIT's Department of Economics), the RAND Corporation, and the Institute for Humane Studies (through the Claude Lambe Fellowship) for their generous financial support of my research. Additionally, the people at IHS (Jeremy Shearmur and Walter Grinder) and the RAND Corporation (Kent Osband and Jim Dertouzous) have been very helpful in their encouragement of my studies. I must also offer a special thanks to Steve Shavell and Harvard Law School's Program in Law and Economics; Steve always welcomed me at the Law School and supported my research in law and economics, even after I discontinued my formal legal studies. Finally, I

would like to thank the Graduate School of Business at the University of Chicago, which by granting me a Visiting Scholar position in the Spring of 1991, afforded me ideal conditions to complete the writing of my thesis.

The research in this thesis has benefited from numerous seminars and audiences. Chapter 1 was presented to audiences at MIT, Yale, Princeton, Chicago, Michigan, Northwestern, Harvard, Stanford, and Columbia. Chapter 2 was given in an earlier form at an MIT theory lunch. Earlier versions of chapter 3 were presented at a RAND conference on defense procurement sponsored by the Policy Analysis and Evaluation Section of the Department of Defense and at a Pew Foundation Conference on Defense Procurement at the National Bureau for Economic Research in Cambridge.

Of course, it goes without saying that any remaining errors in this dissertation are my own.

Biographical Information

Lars Stole was born in St. Louis county on September 22, 1964. He grew up in Bridgeton, Missouri (a suburb of St. Louis), where he attended Pattonville High School from 1979 to 1982. Upon graduation from high school, he went to the University of Illinois at Champaign-Urbana, where he earned bachelor degrees in both economics and political science during the three-year period, 1982-1985, graduating with highest distinction.

From 1985-1986, Stole attended the London School of Economics and Political Science on a Sir Arthur Lewis Studentship, graduating with a Masters of Science in economics in June 1986. He then entered the joint Law and Economics (J.D.-Ph.D.) graduate program at Yale University in the fall of 1986 during which time he studied economics on a National Science Foundation Fellowship. In January 1987, he transferred his economics studies (and NSF fellowship) to the Ph.D. program at the Massachusetts Institute of Technology, where he studied economics through August 1988. During the summer of 1988, Stole was a summer consultant to the RAND Corporation in Santa Monica, California, where he researched theoretical issues in defense procurement which ultimately led to the material contained in chapter 3. In September 1988, Stole began studying law at Harvard University's Law School (J.D. program) as an Olin Fellow, which he discontinued in May 1989. During the summer of 1989, he was a summer associate at the law firm of Perkins-Coie in Washington, D.C., where he specialized in government contract law.

Stole returned to M.I.T. in September 1989, where he has remained until the present Spring of 1991. While at MIT, Stole received an Olin Fellowship from the Olin Foundation and a Claude Lambe Fellowship from the Institute for Humane Studies. Besides the work in this thesis, his research conducted while at MIT includes:

"Licensing and Technology Transfer," in *The Economics of Defense Procurement*, collected papers from the Pew Conference on Defense Procurement (September 22-23, 1989), forthcoming.

"The Economic Effects of Liability Rules on Railroad Employee Accidents: 1890-1970," unpublished manuscript, August 1990.

Solutions to Review Exercises for The Theory of Industrial Organization by Jean Tirole (with Keith Head), MIT Press, 1990.

"Fixed-Equilibrium Rationalizability in Signaling Games," (with Joel Sobel and Inigo Zapater), *Journal of Economic Theory*, 1990, Vol. 52, No. 2, pp.304-331.

Lars Stole is currently a Ph.D. candidate at MIT and a Visiting Scholar at University of Chicago, Graduate School of Business. Beginning in July 1991, Stole will be an Assistant Professor of Business Economics at the GSB. His interests include bicycling, philosophy, film, wine, and jazz.

Contents

Abstract	2
Acknowledgments	5
Biographical Information	8
Introduction	11
1 Mechanism Design Under Common Agency	14
1.1 Introduction	14
1.2 The Model	18
1.2.1 The Contracting Framework	18
1.2.2 The Cooperative Benchmark	21
1.2.3 The Noncooperative Benchmark with Contractual Independence	25
1.3 Incentive Constraints under Common Agency	31
1.3.1 Implementable and Feasible Contracts	31
1.3.2 Strategic Revelation Effects	37
1.4 Analysis of Equilibria with Contract Complements	40
1.5 Analysis of Equilibria with Contract Substitutes	45
1.6 Applications	50
1.6.1 Internal versus External Organization	51
1.6.2 Multiple Regulators	59
1.7 Conclusion and Further Remarks	62
Appendix	63

References	71
2 The Economics of Liquidated Damage Clauses in Contractual Environments with Private Information	73
2.1 Introduction	73
2.2 The Contractual Framework	77
2.3 Contracting Environments	81
2.3.1 The Buyer's Optimal Contract	81
2.3.2 The Seller's Optimal Contract	84
2.3.3 Brokered Contracts	87
2.4 Welfare Implications and Policy Conclusions	90
2.5 Extensions and Conclusions	93
Appendix	95
References	99
3 Information Expropriation and Moral Hazard in Optimal Auctions	100
3.1 Introduction	100
3.2 The Model	103
3.3 The Optimal Contract	107
3.3.1 The First Best	107
3.3.2 Asymmetric Information and Limited Liability	108
3.3.3 The Value of Technology Transfers	111
3.3.4 An Example	113
3.4 Moral Hazard	116
3.4.1 The Problem of Moral Hazard	116
3.4.2 The Model with Moral Hazard	117
3.4.3 The Optimal Contract under Moral Hazard	118
3.5 Conclusions	121
Appendix	123
References	128

Introduction

The 'theory of business' leads a life of obstruction, because theorists do not see the business, and the men of business will not reason out the theories.

– *Walter Bagehot*

A contract is nothing more than an enforceable agreement between two or more parties regarding some activity by which they are all affected. This broad definition is perhaps too encompassing to be descriptive; indeed, contracts govern most economic activity. Even two agents who agree to trade at a price set by a marketplace Walrasian auctioneer are contractual partners.

Contracts, however, are potentially much more interesting than simple marketplace activity in at least two dimensions. First, contractual activity may involve agreements made at dates which are not simultaneous with the economic activity under consideration. For example, investments may be sunk or outside opportunities may arise after contracts are written, but before money is exchanged. Additionally, information regarding an agent's effort or type may appear after the contract is signed. Second, the terms of contracts may be written by one or both parties rather than an anonymous marketplace. For example, sellers may offer prices to buyers which exceed marginal cost and buyers may demand prices below their valuation. In this respect, the study of economic contracts which differ from simple marketplace transactions has considerable interest. This study is largely what economists have in mind when they refer to *contract theory*.

If we wish to construct a "theory of business" that is a reflection of the real world, we are naturally compelled to examine contractual relationships. Buyers frequently interact with sellers in a setting unlike an impersonal marketplace. Relationships within organizations as well as between firms often involve a few rational agents making eco-

conomic arrangements among themselves. Even regulation of business enterprise by the government can be thought of as a contractual relationship between agent and principal.

The three essays in this thesis are examples of the application and extension of contract theory to help understand real economic problems. The first essay considers the contracting problem facing multiple principals, each of whom desires to contract with the same agent. If the agent has private information regarding his gains from the contracting activity and the contracting activities in the principal-agent relationships are substitutable (complementary), the principals will typically extract less (more) information rents in total and induce less (more) productive inefficiency in the contracting equilibrium than if there were a single principal contracting over the same activities. I apply this analysis to various environments, including joint ventures, exclusive-dealing relationships, and regulation between conflicting governmental agencies.

The second essay considers the potential use of liquidated damage clauses under asymmetric information between buyers and sellers in contracts. Courts typically allow parties to stipulate the damages each will pay the other in event of breach, providing that such liquidated damage terms do not greatly exceed actual losses. This restriction acts as a ceiling, however, as courts generally enforce terms that are equal to or below actual losses. Bargaining under asymmetric information can explain this anomaly. Here, the liquidated damage clause serves a dual role, both promoting efficient breach and signaling a party's valuation of trade. I show that it is always optimal for parties to set damages at or below valuation, thereby providing a consistent theory for the courts' asymmetric treatment of contractual damages: When damages are significantly below actual losses, courts may plausibly maintain the presumption that the contract is the result of rational bargaining. But when damages exceed losses, they must consider the likelihood of other elements such as mistake and fraud.

The third essay examines the problem of procurement and second sourcing by the government. When the government procures an item, ideally it would like to pay no more than the minimum possible cost. But in practice the government does not know all of the technological possibilities, even if it could perfectly audit most incurred costs. As a consequence, suppliers who know they have superior technology can earn extra profits –

commonly known as "information rents". One standard way to reduce information rents is to use competitive sealed-bid tenders, which in effect award the contract to the best producer at the second-best bidder's break-even cost. One problem with the preceding method is that only one source may be technically qualified to bid. If government wants more competition, it must "generate" it through licensing or some other form of technology transfer. Conventional cost analysis wisdom asserts that transfer is merited if and only if the second source's direct costs plus the costs of transfer are less than the first source's direct costs. The third essay shows that technology transfer offers an additional potential gain: reduction of information rents. To provide appropriate incentives, technology should perhaps be transferred even when the second source is less efficient than the first. Additionally, when developer moral hazard exists with respect to investments in cost-reducing technology, the optimal auction will make the developer's success in the auction more sensitive to the developing firm's announced costs.

It is my hope that the reader will find that these three essays succeed in furthering our knowledge of both the theory and practice of contracts.

Chapter 1

Mechanism Design Under Common Agency

‘The question is,’ said Humpty Dumpty, ‘which is to be master that’s all.’

– *Lewis Carroll*

1.1 Introduction

Mechanism design has proven to be a fertile area of research for the economist studying the role of information in economic exchange. Since the methodology was first developed by Mirrlees [1971], it has been applied to numerous contexts. Theorists have subsequently extended the use of mechanism design to problems with multi-dimensional type spaces¹, multiple agents², and informed principals.³ But to date, we know very little about the problem of mechanism design with multiple principals and a single agent – what has been termed the problem of *common agency*.⁴

Common agency contracting under adverse selection is ubiquitous. Wherever hidden information and some degree of competition among principals exists for a set of agents,

¹See Rochet[1985], Laffont, Maskin, and Rochet [1987], and McAfee and McMillan [1988].

²See Myerson[1981], Demski, Sappington [1984], Demski, Sappington, and Spiller [1988], and Ma, Moore, and Turnbull [1988].

³See Myerson[1983], and Maskin and Tirole [1990a,1990b].

⁴David Martimort [1991] has independently studied many of the issues in this paper and obtained similar conclusions.

we will generally find an environment where mechanism design under common agency is appropriate. Often the assumption that a single principal completely controls the contracting environment with an agent is not realistic as the following examples illustrate:

- **Multiple regulators.** Several agencies may have authority to promulgate regulations affecting a single agent. To the extent that each regulator (principal) wishes to extract the agent's information rents, an analysis of mechanism design under common agency is appropriate.⁵
- **Common Marketing Agency.** Manufacturers frequently choose to use the same marketing agency for their wares. Such agencies typically have private information about marketing and distribution costs, as well as their effort levels.⁶
- **Price discrimination.** Duopolists selling differentiated products to the same consumers may find it optimal to employ second-degree price discrimination, but must take into account the effect of their rival's nonlinear screening contract.⁷
- **Exclusive Supply Contracts and Joint Ventures.** Firms may decide to form joint ventures with one another to create an exclusive input supplier for members of the venture. In one sense, a joint venture allows firms to coordinate their separate contracts into a single cooperative contract with an agent. In the absence of a joint venture (or alternatively an exclusive supply contract) the firms may non-cooperatively contract with the same agent and fail to take into account the

⁵Related research by Baron [1985] considers a Stackelberg game of regulating a public utility with emission abatement regulation by the EPA (the leader) and rate regulation by a local public utility commission (the follower). This paper extends Baron's approach to a large class of simultaneous contracting games.

⁶This situation was originally considered by Bernheim and Whinston [1985] in an environment of moral hazard. A more general treatment of common agency under conditions of moral hazard is found in Bernheim and Whinston [1986]. Recent work by Villas-Boas [1990] examines the information costs of firms using the same advertising agency, where an agent may tell the "secrets" of one principal to the other. Neither, however, considers adverse selection with common agency. Gal-Or [1989] has also examined a special case of common agency between two principals using the same marketing agent where the utility the agent derives from the relationship with one of the principals is independent of the contract with the other principal. This case is briefly considered in Section 2.3.

⁷Competition with nonlinear tariffs was considered by Oren, et al. [1983], but in a more limited framework where players are restricted to taking the choices of the agent from the rival principal's contract as given.

externalities which they impose on one another. An analysis of common agency illuminates some of the benefits of joint ventures and exclusive supply contracts.⁸

- Franchise Contracts. Franchisors frequently contract with nonexclusive franchisees, such as automobile dealerships, which have contracts with multiple franchisors. The nature of the equilibrium contracts in the nonexclusive environment sheds light on the benefits of exclusive control.
- State and Federal Taxation. Following Mirrlees [1971], an obvious extension of the optimal theory of taxation would consider the effects of two principals (State and Federal revenue departments), each attempting to minimize the distortion introduced by its taxation while maximizing its own objective.

Following the work of Bernheim and Whinston [1986] on common agency under moral hazard, we note that environments with common agency can either be delegated or intrinsic. Under *delegated* common agency, the choice of contractual relationship is delegated to the agent who can choose whether to contract with both, one, or none of the principals. This is a natural setting for examining such phenomena as second degree price discrimination by duopolists, where the consumer ultimately decides from whom to purchase. Alternatively, when common agency is *intrinsic*, the agent's choice is more limited: the agent can choose only between contracting with both principals or contracting with neither. A common example of such a setting is industrial regulation by multiple regulators. The regulated firm's only choice beside regulation is to leave the market and forego profits altogether.

The distinction between these two environments is less important when the contracting activities of the two principals are complementary in terms of the common agent's utility: In any equilibrium where the agent finds it attractive to contract exclusively with either principal, the agent will find it desirable to contract with both. Although this is not the case when the activities are substitutes, we choose to focus on intrinsic common agency

⁸Related models which have examined organizational and market structures from a common agency perspective with moral hazard are Braverman and Stiglitz [1982], which considers sharecroppers responsible to both landlords and creditors, and Stiglitz [1985], which considers corporate managers as agents to both stockholders and corporate creditors.

as a first step toward a more general theory on common agency under adverse selection. Nonetheless, as the applications in this paper demonstrate, a large set of interesting economic questions are addressable within this class of models.

The main focus of this paper is twofold. First, we develop techniques for studying common agency contracts with mechanism design. Second, using these new tools, we consider some of the economic ramifications of a common agency setting. Section 2 of this paper introduces a general model of contracting under common agency, and proceeds by characterizing the contracts for two benchmarks: the cooperative (or single principal) solution and the case of contractual independence (where the agent's marginal utility derived from the contract with one principal is unaffected by the contract with the other).

Two fundamental problems are encountered when one attempts to apply traditional mechanism design tools to common agency problems in absence of contractual independence. First, the simple characterization of incentive compatibility and participation constraints used in single principal contracts is no longer available. Instead, we find a more complicated analog in our two-principal setting when we consider common agency implementability in Section 3. With two principals, each of whom observes only the report meant for her, we require more than that the agent finds it incentive compatible to report truthfully to principal i given he reports truthfully to principal j : It must also be the case that lying to both principals (with perhaps differing reports) is not beneficial to the agent. A significant contribution of this research is to explicitly characterize the set of commonly implementable contracts. Second, when searching for a Nash equilibrium in contracts among principals, one cannot invoke the revelation principle without exercising care. Each principal will typically find it rational to attempt to induce the agent to report falsely to a rival and thereby extract a larger share of the agent's information rents. Of course in equilibrium, all contracts are incentive compatible so that such attempts are useless, but their possibility imposes constraints on the set of equilibrium contracts. This problem is also taken up in Section 3.

Section 4 analyzes the set of pure-strategy differentiable Nash equilibria in the contract game for the cases of contract complements. Section 5 analogously considers equilibria with contract substitutes. We find that the presence of common agency results in each

principal creating a contractual externality. When the contracting activities are complementary, equilibria in the simultaneous contracting game have each principal introducing too much distortion in an effort to extract rents from the agent. With substitutes, the reverse typically occurs and too little distortion is introduced from each principal's point of view. The results are in accord with our notions of Nash equilibria in prices between competing duopolists in a differentiated product market. When the goods for sale are complements, each duopolist prices excessively relative to the monopoly solution; when the goods are substitutes, each duopolist sets prices closer to marginal cost, introducing a smaller distortion. In Section 6 several applications of common agency contracting in environments of adverse selection are presented as a motivation to the preceding analysis. Section 7 concludes.

1.2 The Model

1.2.1 The Contracting Framework

For simplicity we consider a contracting environment with two principals, $i = 1, 2$, and one agent. Although our model is quite general, for exposition we take each principal i as a potential purchaser of some good, x_i , which the agent produces. The agent has private information, or *type*, θ in some compact set Θ , which we take to be the interval $\Theta \equiv [\underline{\theta}, \bar{\theta}]$. Furthermore, it is common knowledge among the principals that θ is distributed according to the differentiable density function $f(\theta)$, where $f(\theta) > 0, \forall \theta \in \Theta$, with corresponding cumulative distribution function $F(\theta)$, and with $\frac{1-F}{f}$ nonincreasing in θ . Without loss of generality, we consider direct revelation mechanisms in which the agent announces his type to each principal separately, although as indicated care must be taken in this regard when considering deviations by each principal from the equilibrium.

We assume that each principal observes only the report meant for her, and denote the reports for each principal as $\hat{\theta}_1$ and $\hat{\theta}_2$, respectively. Various motivations exist to justify this approach. First, antitrust laws might deal harshly with collusive activities to coordinate contracts and reports from the agent, particularly given our results in Section

5 regarding the potential anticompetitive effects of such coordination. Second, even if principals could jointly observe the agent's report, the possibility of secret side contracts between each principal and the agent before the agent's type is announced may render such joint observations useless.⁹ Finally, at least in the regulatory context, it may be legally impossible for one agency to contract on the decision variable of another, even though it may be publicly observed (e.g., the local public utility commission cannot make allowed rates of return an arbitrary function of pollution abatement and the EPA cannot choose levels of allowable pollution as a function of local rate making).

Each principal chooses an *allocation* or contract, $y_i(\cdot)$, which consists of a *decision*, $x_i(\cdot)$, that belongs to a compact, convex, nonempty subset $\mathcal{X} \subset \mathfrak{R}_+$, and a monetary transfer, $t_i(\cdot)$, paid by the principal to the agent: $y_i(\hat{\theta}_i) = \{x_i(\hat{\theta}_i), t_i(\hat{\theta}_i)\}$. We suppose the decision choice of each principal's contract is one-dimensional to simplify the analysis although, as in Guesnerie and Laffont [1984], it is possible to generalize the results to choices over vectors of decisions.

The principals have von Neumann-Morgenstern utility functions that are given by $V^i(x_1, x_2, t_i)$, $i = 1, 2$, which are thrice continuously differentiable, decreasing in t_i , and have partial derivatives up to the third order which are uniformly bounded on any given compact subset of $\mathcal{X}^2 \times \mathfrak{R}_+$. Initially, we let V^i depend upon x_j as in the case where each principal i buys inputs x_i from the agent and sells them in the same downstream product market.

We have chosen to model each principal's utility as a function only of the two contract variables and the transfer to the agent. The agent's type does not affect the principal's welfare. It is straightforward to make each principal's utility a function of θ as well as x_1 and x_2 , although the assumptions used in this paper must be modified to ensure concavity in the principal's problem and monotonicity in the resulting menus of allocations. Such an extension would be appropriate, for example, in the multiple regulators context. In such circumstances, each regulator may place some weight on the agent's welfare (e.g., a

⁹If, however, the side contracts are negotiated under asymmetric information, a role may nonetheless exist for common contracts. See the work of Caillaud, Jullien, and Picard [1990], which shows in a multi-principal and multi-agent framework that if secret contracts are feasible, initial contracts may be useful when asymmetric information exists during side contract negotiation.

public utility's profits may have a positive weight of less than one attached to it), which renders principal i 's payoff a function of x_1, x_2, θ , and t_j as well. Nonetheless, we make the simplifying assumption for ease in exposition. Because each principal's utility depends upon both x_1 and x_2 , the contract between the agent and one of the principals will directly affect the well being of the other principal. More interestingly, to the extent that $U_{x_1 x_2} \neq 0$, one principal's contract will affect the agent's marginal utility, and therefore indirectly affect the cost of contracting with the other principal. Later in this paper we will make a further simplification that each V^i is independent of x_j in order to focus on this second affect.

We assume the agent has a von Neumann-Morgenstern utility function given by

$$U(x_1, x_2, t_1 + t_2, \theta),$$

which is also thrice continuously differentiable, strictly increasing in aggregate transfers, $t_1 + t_2$, and has uniformly bounded partial derivatives up to the third order on any given compact subset of $\mathcal{X}^2 \times \mathfrak{R}$. We also suppose there are no fixed costs of production by the agent: $U(0, 0, 0, \theta) = 0$.

We normalize the agent's outside opportunities to zero and assume that the principals have all of the bargaining power and simultaneously offer take-it-or-leave-it contracts. Because we analyze intrinsic agency, we suppose that the agent is forced either to accept both contracts or to refuse to contract with both principals.

Given a contract pair, $\{y(\hat{\theta})\} = \{y_1(\hat{\theta}_1), y_2(\hat{\theta}_2)\}_{\hat{\theta}_i \in \Theta, i=1,2}$, we can represent an agent's indirect utility as a function of reports and type by

$$U(\hat{\theta}_1, \hat{\theta}_2, \theta) \equiv U(x_1(\hat{\theta}_1), x_2(\hat{\theta}_2), t_1(\hat{\theta}_1) + t_2(\hat{\theta}_2), \theta),$$

which we will frequently use when no confusion should result. Additionally, subscripts denote partial derivatives with respect to direct arguments and primes denote derivatives with respect to a single argument at all points where such derivatives exist.

1.2.2 The Cooperative Benchmark

As a comparison, we initially consider the situation where both principals choose contracts that depend upon a single report by the agent and that maximize their joint utilities.¹⁰ [The reader familiar with the theory of mechanism design may wish to skip to Section 2.3.] Alternatively, we can think of the situation as one of a single principal that contracts over both activities of the agent. As a consequence, we can restrict ourselves to a simple mechanisms $y(\hat{\theta}) = \{t(\hat{\theta}), x_1(\hat{\theta}), x_2(\hat{\theta})\}$, where $\hat{\theta}$ is the single report by the agent. Given an allocation, we may denote the agent's utility as a function of type and report by $U(\hat{\theta}, \theta) \equiv U(x_1(\hat{\theta}), x_2(\hat{\theta}), t(\hat{\theta}), \theta)$.

Definition 1 *A decision function, $x : \Theta \mapsto \mathcal{X}^2$, is implementable if there exists a transfer function $t(\cdot)$ such that the contract satisfies the incentive compatibility (IC) constraint:*

$$U(x_1(\theta), x_2(\theta), t(\theta), \theta) \geq U(x_1(\hat{\theta}), x_2(\hat{\theta}), t(\hat{\theta}), \theta), \quad \forall (\theta, \hat{\theta}) \in \Theta^2.$$

A contract is feasible if the decision function is implementable, and the transfers additionally satisfy the participation (or individual rationality) constraint:

$$U(x_1(\theta), x_2(\theta), t(\theta), \theta) \geq 0, \quad \forall \theta \in \Theta.$$

Throughout this paper we will restrict ourselves to continuous decision functions which have piecewise continuous first derivatives (i.e., are *piecewise C^1*). Following the methodology in Mirrlees [1971] we may characterize the set of feasible mechanisms in the following two theorems.¹¹ Although the results of Theorems 1 and 2 are standard, we present them in the Appendix for completeness and comparison with the proofs used in characterizing implementability and feasibility under common agency.

¹⁰In the general case where U is not linear in transfers, we may look for a Pareto optimum such that $\lambda V^1 + (1 - \lambda)V^2$ is maximized for some weight, λ . When U is quasi-linear we may consider the simple sum of the principals' payoffs. Here we focus on the latter.

¹¹This section closely follows the development in Guesnerie and Laffont [1984]. For another exposition, combined with a more recent review of the literature, see Fudenberg and Tirole [1991, chapter 7].

Theorem 1 (Necessary Conditions.) *A piecewise C^1 decision function is implementable only if*

$$U_t(x_1, x_2, t, \theta)t'(\theta) = - \sum_{i=1}^2 U_{x_i}(x_1, x_2, t, \theta)x_i'(\theta), \quad (1.1)$$

and

$$\frac{\partial}{\partial \theta} \left(\frac{U_{x_1}(x_1, x_2, t, \theta)}{U_t(x_1, x_2, t, \theta)} \right) x_1'(\theta) + \frac{\partial}{\partial \theta} \left(\frac{U_{x_2}(x_1, x_2, t, \theta)}{U_t(x_1, x_2, t, \theta)} \right) x_2'(\theta) \geq 0, \quad (1.2)$$

for any θ such that $x_i = x_i(\theta)$, $t = t(\theta)$ are differentiable at θ , which is the case except at a finite number of points. In addition, an allocation is feasible only if

$$U(x_1(\underline{\theta}), x_2(\underline{\theta}), t(\underline{\theta}), \underline{\theta}) \geq 0. \quad (1.3)$$

Before proceeding with the sufficiency theorem, we make two assumptions.

Assumption 1 *Constant sign of the marginal rate of substitution. On the relevant domain of x_1 , x_2 , t , and θ , $\frac{\partial}{\partial \theta} \left(\frac{U_{x_i}(x_1, x_2, t, \theta)}{U_t(x_1, x_2, t, \theta)} \right) > 0$, $i = 1, 2$. Additionally, the agent's utility increases in θ : $U_\theta(x_1, x_2, t, \theta) > 0, \forall x_1, x_2, t, \theta$.*

Assumption 2 *Boundary behavior of $U(\cdot)$. For any $(x_1, x_2, t, \theta) \in \mathcal{X}^2 \times \mathfrak{R} \times \Theta$, there exists a $K > 0$ such that*

$$\left\| \sum_{i=1}^2 \left[\frac{U_{x_i}(x_1(\theta), x_2(\theta), t, \theta)}{U_t(x_1(\theta), x_2(\theta), t, \theta)} - \frac{U_{x_i}(x_1(\theta), x_2(\theta), t', \theta)}{U_t(x_1(\theta), x_2(\theta), t', \theta)} \right] \frac{dx_i(\theta)}{d\hat{\theta}} \right\| \leq K \|t - t'\|,$$

uniformly in x_1 , x_2 , and θ , where $\|\varphi\| = \sup_{\theta \in \Theta} |\varphi(\theta)|$.

Assumption A.1 is the well known Spence-Mirrlees single-crossing condition; this partial derivative exists because U is C^2 and strictly increasing in t . Without loss of generality, we assume the signs are positive. The condition that the agent's utility increases in θ is natural in most economic environments where the marginal rate of substitution between activity and transfer is positive. We take A.1 as given throughout this paper.

Assumption A.2 is a Lipschitz condition which assures us that the marginal rates of substitution between decisions and transfers do not increase too fast when the transfer increases. With preferences that are linear in transfers, this condition is trivially satisfied. We now state the sufficiency theorem.

Theorem 2 (Sufficient Conditions.) *Given assumptions A.1-A.2, any piecewise C^1 decision profile for which $x_i^!(\theta) \geq 0, \forall \theta \in \Theta, i = 1, 2$, is implementable by a transfer function satisfying (1.1). Furthermore, given that a piecewise C^1 allocation satisfies condition (1.3), the allocation is also feasible.*

The traditional approach to mechanism design takes (1.1) and (1.3) above and chooses a mechanism which maximizes the principal's utility. It is then checked that the resulting mechanism is monotone. In the event that it is not, an algorithm such as that in Guesnerie and Laffont [1984] is employed which monotonizes the decision functions in an optimal manner. In the present case of cooperative contracts, we may proceed accordingly. First, however, for tractability in the principals' optimization problem, we make additional assumptions regarding the contracting environment.

Assumption 3 (a) *Agent's preferences are quasi-linear: $U(x_1, x_2, t, \theta) = U(x_1, x_2, \theta) + t$.*

(b) *Principals' preferences are quasi-linear: $V^i(x_1, x_2, t_i) = \mathcal{V}^i(x_1, x_2) - t_i$.*

(c) *The range of allowable decision functions, \mathcal{X} , is the interval $[0, \bar{x}]$, where (\bar{x}, \bar{x}) is greater than any $(x_1, x_2) \in \arg \max_{x_1, x_2} \{U(x_1, x_2, \bar{\theta}) + \mathcal{V}^i(x_1, x_2)\}$, for $i = 1, 2$ and greater than any $(x_1, x_2) \in \arg \max_{x_1, x_2} \{U(x_1, x_2, \bar{\theta}) + \mathcal{V}^1(x_1, x_2) + \mathcal{V}^2(x_1, x_2)\}$.*

Assumption 4 Concavity and monotonicity.

(a) *The following function (the principals' virtual surplus) is globally strictly concave in x_1 and x_2 , and for all θ attains an interior maximum over \mathcal{X}^2 :*

$$\mathcal{V}^1(x_1, x_2) + \mathcal{V}^2(x_1, x_2) + U(x_1, x_2, \theta) - \frac{1-F(\theta)}{f(\theta)} \mathcal{U}_\theta(x_1, x_2, \theta);$$

additionally, $\mathcal{U}_{\theta\theta}(x_1, x_2, \theta) \leq 0$.

(b) *For $i = 1, 2$, and for any x_1, x_2, θ ,*

$$\left[\mathcal{V}_{x_1 x_2}^1 + \mathcal{V}_{x_1 x_2}^2 + \mathcal{U}_{x_1 x_2} - \frac{1-F(\theta)}{f(\theta)} \mathcal{U}_{x_1 x_2 \theta} \right] \left[\mathcal{U}_{x_j \theta} \left(1 - \frac{d}{d\theta} \left(\frac{1-F(\theta)}{f(\theta)} \right) \right) - \frac{1-F(\theta)}{f(\theta)} \mathcal{U}_{x_j \theta \theta} \right] -$$

$$\left[\mathcal{V}_{x_j x_j}^1 + \mathcal{V}_{x_j x_j}^2 + \mathcal{U}_{x_j x_j} - \frac{1-F(\theta)}{f(\theta)} \mathcal{U}_{x_j x_j \theta} \right] \left[\mathcal{U}_{x_i \theta} \left(1 - \frac{d}{d\theta} \left(\frac{1-F(\theta)}{f(\theta)} \right) \right) - \frac{1-F(\theta)}{f(\theta)} \mathcal{U}_{x_i \theta \theta} \right] \geq 0.$$

Although assumption A.3(a)-(b) is strong, it allows us to get to the heart of the issues of adverse selection under common agency without introducing additional technical

assumptions. Nonetheless, it should be clear to the reader how one proceeds when preferences are not quasi-linear. In our context of two principals buying products from a single supplier (agent), \mathcal{U} represents the costs of production and is negative, while t represents revenues from the principals. A.3(c) additionally requires that the principals are not specifically prevented from implementing the first-best level of activity.

A.4(a) assumes that the principals' incomplete-information problem is well-behaved. This assumption is met whenever the full-information optimum is globally strictly concave (as is the case in many economic problems) and the uncertainty of θ is relatively small. In the absence of A.4(a), it is possible that corner solutions as well as random schemes may be desirable. The condition that $\mathcal{U}_{\theta\theta} \leq 0$ ensures that at the optimum, the expression in A.4(a) is increasing in θ .

Unless a particular economic environment is considered, assumption A.4(b) is not naturally satisfied. A.4(b) (in combination with A.1, A.3, and A.4(a)) requires that the unconstrained solution to the principals' incomplete information problem have increasing decision functions. This simplifies our task considerably, as we do not have to consider such issues as bunching. Sufficient (but not necessary) conditions for A.4(b) to hold are $\mathcal{U}_{x_i\theta\theta} \leq 0$ and $\mathcal{U}_{x_1x_2\theta} \geq 0$. Section 6 provides motivating economic applications that satisfy A.4(b).

Given the additional assumptions A.3-A.4, we can now state the solution to the principals' cooperative contracting problem.

Proposition 1 *Given assumptions A.3 and A.4, the contract which maximizes the sum of the principals' utilities has decision functions which satisfy $\forall \theta \in [\theta_i^*, \bar{\theta}]$, $i = 1, 2$*

$$\mathcal{V}_{x_i}^1(x_1, x_2) + \mathcal{V}_{x_i}^2(x_1, x_2) + \mathcal{U}_{x_i}(x_1, x_2, \theta) = \frac{1 - F(\theta)}{f(\theta)} \mathcal{U}_{x_i\theta}(x_1, x_2, \theta), \quad (1.4)$$

and $\forall \theta \in [\underline{\theta}, \theta_i^*)$, $x_i(\theta) = 0$, where θ_i^* is defined by

$$\begin{aligned} &\mathcal{V}^1(x_1(\theta_i^*), x_2(\theta_i^*)) + \mathcal{V}^2(x_1(\theta_i^*), x_2(\theta_i^*)) + \mathcal{U}(x_1(\theta_i^*), x_2(\theta_i^*), \theta_i^*) \\ &\quad - \frac{1 - F(\theta)}{f(\theta)} \mathcal{U}_\theta(x_1(\theta_i^*), x_2(\theta_i^*), \theta_i^*) = 0, \end{aligned}$$

if the resulting $\theta_i^* \geq \underline{\theta}$, and $\theta_i^* = \underline{\theta}$ otherwise. Moreover, the transfer function in the optimal contract satisfies $\forall \theta \in \Theta$.

$$t(\theta) = \int_{\underline{\theta}}^{\theta} \mathcal{U}_{\theta}(x_1(s), x_2(s), s) ds - \mathcal{U}(x_1(\underline{\theta}), x_2(\underline{\theta}), \underline{\theta}). \quad (1.5)$$

The proof of the proposition is standard and provided in the appendix. Proposition 1 indicates that the contracted levels of x_i are below the efficient level for all $\theta < \bar{\theta}$. The intuition behind the result is straightforward. The principals contract for levels of x_i for a given θ such that the marginal expected efficiency gain from raising the level of x_i , i.e. $(\mathcal{V}_{x_i}^1 + \mathcal{V}_{x_i}^2 + \mathcal{U}_{x_i})f(\theta)$, is equal to the marginal loss of rents which must be given to agents with types better than θ to induce incentive compatibility, i.e. $\mathcal{U}_{x_i, \theta}[1 - F(\theta)]$. Of course, when the principals have unaligned preferences (i.e., neither principal cares about maximizing the joint surplus) and the contracts are chosen noncooperatively, this result is fundamentally altered.

In the noncooperative contracting game in which the principals have different preferences for contracting activities, the presence of externalities alters the result in Proposition 1. Two channels exist for the transmission of externalities. First, when \mathcal{V}^j depends on x_i , principal i will not take into account \mathcal{V}^j when maximizing her payoffs and may choose x_i inefficiently from the point of view of maximizing joint surplus. We examine this effect in the following section. The second channel which exists even if \mathcal{V}^j is independent of x_i , is both more interesting and more subtle. To the extent that $\mathcal{U}_{x_1 x_2} \neq 0$, the contract of one principal may change the marginal disutility to the agent from the other principal's contracting activity, thereby affecting the equilibrium contracts offered by each principal. The examination of this second channel is undertaken in the remainder of this paper.

1.2.3 The Noncooperative Benchmark with Contractual Independence

We now depart from the earlier analysis where we assumed that the two principals could coordinate contracts with the agent, and where each principal learned of both reports. Instead we suppose a common agency environment where each principal may condition her contract only upon the report meant for her that is sent by the agent. Each principal's

mechanism, $y_i(\cdot) = \{x_i(\cdot), t_i(\cdot)\}$, is a function only of $\hat{\theta}_i$. Such a representation is equivalent to the nonlinear tariff contract where $t_i = t_i(x_i)$, and t_i is independent of x_j .

Under full-information, a set of equilibrium contracts which maximizes the principals' joint surplus exists where each principal makes the agent the residual claimant for her profit, thereby internalizing the externalities the principals would otherwise impose upon one another. When information is private, we must again address the issue of incentive compatibility.

As before, given a pair of contracts and our assumption of quasi-linear payoffs, we can denote the utility of an agent with type θ who makes reports $\hat{\theta}_i$ to principal i as

$$U(\hat{\theta}_1, \hat{\theta}_2, \theta) \equiv \mathcal{U}(x_1(\hat{\theta}_1), x_2(\hat{\theta}_2), \theta) + t_1(\hat{\theta}_1) + t_2(\hat{\theta}_2).$$

With this definition, we can define incentive compatibility for the common agency contracting environment.

Definition 2 *A pair of decision functions, $\{x_1(\cdot), x_2(\cdot)\}$, where $x_i : \Theta \mapsto \mathcal{X}$, is commonly implementable if there exists a transfer function $t_i(\cdot) : \Theta \mapsto \mathfrak{R}$ for each principal such that the pair of contracts satisfies the common incentive compatibility (CIC) constraint:*

$$U(\theta, \theta, \theta) \geq U(\hat{\theta}_1, \hat{\theta}_2, \theta), \quad \forall (\hat{\theta}_1, \hat{\theta}_2, \theta) \in \Theta^3.$$

A pair of contracts, $y : \Theta^2 \mapsto \mathcal{X}^2 \times \mathfrak{R}^2$, is commonly feasible if the decision functions are implementable, and the transfers satisfy the participation (or individual rationality) constraint:

$$U(\theta, \theta, \theta) \geq 0, \quad \forall \theta \in \Theta.$$

For completeness we consider the simple case of contractual independence in agent's utility as a benchmark. When the agent's utility from contracting with one principal is independent of the contracting activity with the other (i.e., $\mathcal{U}_{x_1 x_2} = 0$ for all x_1, x_2, θ), the equilibrium of the common agency contracting game is readily calculated. With contractual independence, we abstract away from concerns imposed by global incentive compatibility which manifest themselves whenever the agent can make two different

reports – one to each principal. This benchmark, however, is intriguing as it highlights the strategic interactions which result from our assumption of intrinsic agency and the contracting requirement of individual rationality.

Because the activities are independent from the agent's viewpoint when $\mathcal{U}_{x_1 x_2} = 0$ and A.3(a) holds, Theorems 1 and 2 still apply with only slight modifications in their statements.

Theorem 1' (Necessary Conditions.) *Suppose $\mathcal{U}_{x_1 x_2} = 0$. A piecewise C^1 decision function is implementable only if*

$$t'_i(\theta) = -\mathcal{U}_{x_i}(x_1, x_2, \theta)x'_i(\theta),$$

and $x'_i(\theta) \geq 0$, for any θ such that $x_i = x_i(\theta)$, $t = t_i(\theta)$ are differentiable at θ , which is the case except at a finite number of points. In addition, an allocation is feasible only if

$$\mathcal{U}(x_1(\underline{\theta}), x_2(\underline{\theta}), \underline{\theta}) + t_1(\underline{\theta}) + t_2(\underline{\theta}) \geq 0.$$

Theorem 2' (Sufficient Conditions.) *Suppose that $\mathcal{U}_{x_1 x_2} = 0$. Any piecewise C^1 decision function, x_i , for which $x'_i(\theta) \geq 0$, is implementable by a transfer function, $t_i(\cdot)$, satisfying the differential equation in Theorem 1' above. Furthermore, given that a piecewise C^1 allocation satisfies condition the necessary individual rationality condition in Theorem 1', the allocation is also feasible.*

The proofs follow those from Theorems 1 and 2. Note, however, that the necessary individual rationality condition in Theorem 1' requires principal i 's contract to satisfy a *global* participation constraint. This is an artifact of our intrinsic agency framework. With delegated agency, this condition would be replaced with the participation constraint specific to principal i : $\mathcal{U}(x_1(\theta), x_2(\theta), \theta) + t_i(\theta) \geq \mathcal{U}(0, x_2(\theta), \theta)$. With intrinsic agency, however, we have the possibility that one principal may pay less than her implicit share for the agent's production. This will have an affect on the characterization of equilibrium

contracts.

To proceed with our examination of the contractual independence equilibrium, we modify A.4 as follows:

Assumption 4' Concavity. (a) In addition to A.4(a) holding, the following function (principal i 's virtual surplus) is globally strictly concave in x_i , and for all x_j and θ attains an interior maximum over \mathcal{X} :

$$\mathcal{V}^i(x_1, x_2) + \mathcal{U}(x_1, x_2, \theta) - \frac{1-F(\theta)}{f(\theta)}\mathcal{U}_\theta(x_1, x_2, \theta).$$

(b) For all $x_1, x_2, \theta, i = 1, 2, j \neq i$,

$$\begin{aligned} \Phi^j(x_1, x_2, \theta) &\equiv \psi^j(x_1, x_2, \theta) \left[\mathcal{V}_{x_i x_i}^i(x_1, x_2) + \mathcal{U}_{x_i x_i}(x_1, x_2, \theta) - \frac{1-F(\theta)}{f(\theta)}\mathcal{U}_{x_i x_i \theta}(x_1, x_2, \theta) \right] \\ &- \psi^i(x_1, x_2, \theta)\mathcal{V}_{x_1 x_2}^j(x_1, x_2) \geq 0, \end{aligned}$$

where

$$\psi^i(x_1, x_2, \theta) \equiv \frac{1-F(\theta)}{f(\theta)}\mathcal{U}_{x_i \theta \theta}(x_1, x_2, \theta) - \left[1 - \frac{d}{d\theta} \left(\frac{1-F(\theta)}{f(\theta)} \right) \right] \mathcal{U}_{x_i \theta}(x_1, x_2, \theta).$$

(c) For all x_1, x_2, θ and $i = 1, 2, j \neq i$,

$$\left(\mathcal{V}_{x_j}^i - \frac{1-F}{f}\mathcal{U}_{\theta x_j} \right) \frac{\Phi^j}{\det \Omega} + \left[1 - \frac{d}{d\theta} \left(\frac{1-F}{f} \right) \right] \mathcal{U}_\theta - \frac{1-F}{f}\mathcal{U}_{\theta \theta} \geq 0,$$

where Ω is the Hessian of the expression in A.4'(a).

A.4 has been modified in three ways in order to deal with the strategic interactions induced by the externalities inherent in the principal's payoffs. First, concavity is assumed over an individual principal's objective function. Second, in A.4'(b) conditions related to concavity have been assumed to ensure that $x'_i(\theta) \geq 0$. These latter conditions are satisfied if, for example, $\mathcal{U}_{x_i \theta \theta} \leq 0$ and $\mathcal{V}_{x_1 x_2}^j$ is not too negative relative to $\mathcal{V}_{x_i x_i}^i + \mathcal{U}_{x_i x_i}$; in this sense, A.4'(b) is akin to sufficient concavity of the full information collective surplus.

Third, A.4'(c) effectively requires that principal i 's virtual profits be nondecreasing in θ . The condition is satisfied, for example, whenever $U_{\theta\theta} \leq 0$ and any negative externality born by principal i from x_j is small on the margin compared to the information rents paid to the agent. With A.4' satisfied, we can now state our result.

Proposition 2 *Given assumption A.4' and $U_{x_1 x_2} = 0, \forall x_1, x_2, \theta$, any pure-strategy Nash equilibrium in the simultaneous contracting game satisfies $\forall \theta \in [\theta_i^*, \bar{\theta}]$*

$$V_{x_i}^i(x_1, x_2) + U_{x_i}(x_1, x_2, v_j) = \frac{1 - F(\theta)}{f(\theta)} U_{x, \theta}(x_1, x_2, \theta), \quad i = 1, 2, \quad (1.6)$$

and for all $\theta \in [\underline{\theta}, \theta_i^*)$, $x_i(\theta) = 0$, where θ_i^* is defined by

$$V^i(x_1(\theta_i^*), x_2(\theta_i^*)) + U(x_1(\theta_i^*), x_2(\theta_i^*), v_i^*) - \frac{1 - F(\theta)}{f(\theta)} U_{\theta}(x_1(\theta_i^*), x_2(\theta_i^*), \theta_i^*) + t_j(\theta_i^*) = 0,$$

if the resulting $\theta_i^* \geq \underline{\theta}$, and $\theta_i^* = \underline{\theta}$ otherwise. Moreover, the transfer function in the optimal contract satisfies $\forall \theta \in [\theta_i^*, \bar{\theta}]$

$$t_i(\theta) = \int_{\theta_i^*}^{\theta} U_{x_i}(x_1(s), x_2(s), s) x_i'(s) ds - U(x_1(\theta_i^*), x_2(\theta_i^*), \theta_i^*) - t_j(\theta_i^*), \quad (1.7)$$

and $t_i(\theta) = 0$ for all $\theta \in [\underline{\theta}, \theta_i^*)$.

The proof is analogous to that of Proposition 1 and is discussed in the Appendix. Two principals simultaneously maximize their payoffs. Depending upon the relationship between the principals' payoffs, the resulting contracts can either require greater or lesser contracting activity. If $V_{x_j}^i < 0$, a principal's contract introduces a negative externality, and production is greater under common agency than under the cooperative contract. The opposite conclusion holds for $V_{x_j}^i > 0$. This result is related to the work of Gal-Or [1989], who argues that common agency may impose a cost on the principals in a common marketing relationship. Increased sales of one principal's product by a marketing agent hurts the second principal through reductions in demand.

An additional difference with the result in Proposition 1 involves the nature of the cutoff types, θ_i^* . Because intrinsic agency requires that each principal's contract satisfy

global participation, it is possible that multiple equilibria exist. Supposing that principal i pays only a small fraction of $\mathcal{U}(x_1(\underline{\theta}), x_2(\underline{\theta}), \underline{\theta})$, principal j may find it worthwhile to contract only with $\theta \geq \theta_j^*$. That is, it may be too costly for principal j to pay the difference in order to satisfy the global participation constraint for $\theta < \hat{\theta}_j^*$. Consequently, the equilibrium share, α_i , of $\mathcal{U}(x_1(\underline{\theta}), x_2(\underline{\theta}), \underline{\theta})$ that principal i pays may be required to lie inside a subinterval of $[0,1]$ in order for all types to be contracted. We can say more about the nature of such shares as the following corollary suggests.

Corollary 1 *Suppose each principal's contribution to the joint surplus is positive at $\underline{\theta}$ for a pair of decision functions x_1, x_2 which satisfy (1.6) above: i.e.,*

$$\mathcal{V}^i(x_1(\underline{\theta}), x_2(\underline{\theta})) + \mathcal{U}(x_i(\underline{\theta}), 0, \underline{\theta}) - \frac{1 - F(\underline{\theta})}{f(\underline{\theta})} \mathcal{U}_\theta(x_i(\underline{\theta}), 0, \underline{\theta}) \geq 0.$$

Then there exists t_1, t_2 such that x_1, x_2 is a Nash equilibrium and $\theta_i^ = \underline{\theta}$ for $i = 1, 2$.*

Proof: For $\theta_i^* = \underline{\theta}$, it must be that

$$\mathcal{V}^i(x_1(\underline{\theta}), x_2(\underline{\theta})) + \alpha_i \mathcal{U}(x_1(\underline{\theta}), x_2(\underline{\theta}), \underline{\theta}) - \frac{1 - F(\underline{\theta})}{f(\underline{\theta})} \mathcal{U}_\theta(x_1(\underline{\theta}), x_2(\underline{\theta}), \underline{\theta}) \geq 0,$$

for $i = 1, 2$. Since $\mathcal{U}(0, 0, \underline{\theta}) = 0$ and $\mathcal{U}_{x_1 x_2} = 0$, $\mathcal{U}(x_1(\underline{\theta}), x_2(\underline{\theta}), \underline{\theta}) = \mathcal{U}(x_1(\underline{\theta}), 0, \underline{\theta}) + \mathcal{U}(0, x_2(\underline{\theta}), \underline{\theta})$. Setting $\alpha_i = \mathcal{U}(x_i(\underline{\theta}), 0, \underline{\theta}) / \mathcal{U}(x_1(\underline{\theta}), x_2(\underline{\theta}), \underline{\theta})$, and

$$t_i(\theta) = \int_{\underline{\theta}}^{\theta} \mathcal{U}_{x_i}(x_1(s), x_2(s), s) x_i'(s) ds - \alpha_i \mathcal{U}(x_1(\underline{\theta}), x_2(\underline{\theta}), \underline{\theta}),$$

we satisfy the required condition for $i = 1, 2$. □

The proof uses an α_i set equal to the ratio of principal i 's production cost to the total production cost so as to obtain full contacting by both principals. In fact, an interval for α_i defined by

$$\left[\frac{\mathcal{U}(x_1, x_2, \theta) + \mathcal{V}^j(x_1, x_2) - \frac{1-F(\theta)}{f(\theta)} \mathcal{U}_\theta(x_1, x_2, \theta)}{\mathcal{U}(x_1, x_2, \theta)}, \frac{\frac{1-F(\theta)}{f(\theta)} \mathcal{U}_\theta(x_1, x_2, \theta) - \mathcal{V}^i(x_1, x_2)}{-\mathcal{U}(x_1, x_2, \theta)} \right]$$

exists at each θ such that, for all α_i contained in the interval, all types of agent greater than θ are contracted with by the principals. Any α_i which lies in the interval defined at $\underline{\theta}$ will support the Nash equilibrium given by (1.6) and $\theta_i^* = \underline{\theta}$.

In order to more fully understand the ramifications of common agency in contexts of adverse selection, we now focus our attention to the more subtle problem of non-independent contracting activities.

1.3 Incentive Constraints under Common Agency

1.3.1 Implementable and Feasible Contracts

As in Section 2.3, we suppose a common agency environment where each principal may condition her contract only upon the report meant for her: a principal's mechanism, $\{y_i(\hat{\theta}_i)\} = \{x_i(\hat{\theta}_i), t_i(\hat{\theta}_i)\}_{\hat{\theta}_i \in \Theta}$, may depend only upon $\hat{\theta}_i$. In this Section we characterize a set of necessary and sufficient conditions for common incentive compatibility and participation when $\mathcal{U}_{x_1 x_2} \neq 0$, but for simplicity we assume no externalities between the principals' payoffs (i.e., $\mathcal{V}_{x_j}^i = 0$). Because each principal's contract can only depend upon $\hat{\theta}_i$, the necessary and sufficient conditions will be stronger than in (1.1)-(1.3) above. With conditions similar to (1.1)-(1.3), we can only guarantee that an agent will not make consistent reports, $\hat{\theta}_1 = \hat{\theta}_2$, that differ from θ . Stronger conditions must be satisfied to guarantee in addition that the agent will not gain from making inconsistent lies.

We proceed with two theorems analogous to the necessity and sufficiency theorems presented in Section 2.

Theorem 3 (Necessary Conditions.) *A pair of piecewise C^1 decision functions are commonly implementable only if, for $i = 1, 2$,*

$$U_{\hat{\theta}_i}(\theta, \theta, \theta) = 0, \quad (1.8)$$

$$U_{\hat{\theta}_i, \theta}(\theta, \theta, \theta) + U_{\hat{\theta}_i, \hat{\theta}_2}(\theta, \theta, \theta) \geq 0, \quad (1.9)$$

$$U_{\hat{\theta}_1, \theta}(\theta, \theta, \theta)U_{\hat{\theta}_2, \theta}(\theta, \theta, \theta) + U_{\hat{\theta}_1, \hat{\theta}_2}(\theta, \theta, \theta) \left(U_{\hat{\theta}_1, \theta}(\theta, \theta, \theta) + U_{\hat{\theta}_2, \theta}(\theta, \theta, \theta) \right) \geq 0, \quad (1.10)$$

for any $x_i(\theta)$, $t_i(\theta)$, $\theta \in \Theta$ such that x_i is differentiable at θ . In addition, a pair of piecewise C^1 contracts is commonly feasible only if

$$U(\underline{\theta}, \underline{\theta}, \underline{\theta}) \geq 0. \quad (1.11)$$

Proof: As in Theorem 1, using a Taylor expansion and revealed preference, it can be shown that piecewise C^1 decision functions imply that transfer functions are also piecewise C^1 .

A necessary condition for maximization by the agent is the satisfaction of first-order and local second-order conditions at $\hat{\theta}_1 = \hat{\theta}_2 = \theta$, at all points of differentiability. This implies

$$U_{\hat{\theta}_i}(\theta, \theta, \theta) = 0, \quad i = 1, 2,$$

$$U_{\hat{\theta}_i \hat{\theta}_i}(\theta, \theta, \theta) \leq 0, \quad i = 1, 2,$$

$$U_{\hat{\theta}_1 \hat{\theta}_1}(\theta, \theta, \theta)U_{\hat{\theta}_2 \hat{\theta}_2}(\theta, \theta, \theta) - \left(U_{\hat{\theta}_1 \hat{\theta}_2}(\theta, \theta, \theta) \right)^2 \geq 0,$$

$\forall \theta \in (\underline{\theta}, \bar{\theta})$. The first expression is (1.8) above. Totally differentiating this expression with respect to θ yields $U_{\hat{\theta}_i \hat{\theta}_i}(\theta, \theta, \theta) + U_{\hat{\theta}_i \theta}(\theta, \theta, \theta) + U_{\theta \hat{\theta}_i}(\theta, \theta, \theta) = 0$, $i = 1, 2$, which allows us equivalently to express the local second-order conditions (the second and third expressions above) as (1.9) and (1.10).¹² Finally, feasibility implies (1.11) trivially. \square

Using the implication of quasi-linearity that $\mathcal{U}_t = 1$, we can equivalently state (1.8)-(1.10) in simpler form.

Corollary 2 *A pair of piecewise C^1 decision functions are commonly implementable only if*

$$t'_i(\theta) = -\mathcal{U}_{x_i}(x_1, x_2, \theta)x'_i(\theta), \quad i = 1, 2, \quad (1.12)$$

$$\mathcal{U}_{x_1 x_2}(x_1, x_2, \theta)x'_1(\theta)x'_2(\theta) + \mathcal{U}_{x_i \theta}(x_1, x_2, \theta)x'_i(\theta) \geq 0, \quad i = 1, 2, \quad (1.13)$$

¹²Because x_i is piecewise C^1 , we know that $U_{\hat{\theta}_i \theta}$ exists everywhere but at a finite set of points. Additionally, with A.3, $U_{\hat{\theta}_i \hat{\theta}_i} = \mathcal{U}_{x_1 x_2}(x_1, x_2, \theta)x'_1(\theta)x'_2(\theta)$ which also exists everywhere but at a finite set of points. Thus, a Taylor expansion of $U_{\hat{\theta}_i}$ around θ yields the existence of $U_{\hat{\theta}_i \hat{\theta}_i}$ at all but a finite number of points.

$$\mathcal{U}_{x_1, \theta} \mathcal{U}_{x_2, \theta} x_1'(\theta) x_2'(\theta) + \mathcal{U}_{x_1, x_2} x_1'(\theta) x_2'(\theta) [\mathcal{U}_{x_1, \theta} x_1'(\theta) + \mathcal{U}_{x_2, \theta} x_2'(\theta)] \geq 0. \quad (1.14)$$

for any $x_i(\theta), t_i(\theta), \theta \in \Theta$ such that x_i is differentiable at θ , where the arguments of \mathcal{U} are understood to be $x_1(\theta), x_2(\theta), \theta$.

In what follows, it will be useful to distinguish between two cases of contractual spillovers: contractual complements and substitutes. $\mathcal{U}_{x_1, x_2} > 0$ corresponds to the case where the agent's activities are *contract complements*, while $\mathcal{U}_{x_1, x_2} < 0$ corresponds to the case of *contract substitutes*.

Following Theorem 3, we can say something about the characteristics of commonly implementable contracts.

Corollary 3 *If the contracting activities are complements, a pair of piecewise C^1 decision functions are commonly implementable only if each principal's decision function has a nonnegative derivative at all points of differentiability.*

Proof: Suppose otherwise. Suppose without loss of generality that only x_1 is decreasing over some interval of Θ , while x_2 is nondecreasing. By $\mathcal{U}_{x_1, x_2} > 0$, (1.13) is violated. Suppose instead that each x_i is decreasing over some interval of Θ . (1.9) implies that $\mathcal{U}_{x_1, x_2} x_1' x_2' \geq -\mathcal{U}_{x_1, \theta} x_1'$. (1.14) implies that

$$\mathcal{U}_{x_1, \theta} \mathcal{U}_{x_2, \theta} x_1' x_2' - \mathcal{U}_{x_1, \theta} x_1' (\mathcal{U}_{x_1, \theta} x_1' + \mathcal{U}_{x_2, \theta} x_2') \geq 0,$$

which contradicts our assumption that $\mathcal{U}_{x_i, \theta} > 0$. □

When the contracting activities are substitutes, the analysis is slightly more complicated. The necessary conditions in Theorem 3 are insufficient to prove that both decision functions are monotonically increasing. Instead, it is possible that one schedule may be decreasing if the other is sufficiently increasing. We can only be certain at this point that both functions may not be decreasing over the same interval. We will find in Section 5, however, that under some simplifying conditions on preferences and the distribution of θ both decision functions will be increasing in equilibrium.

The corollary makes clear that in a common agency environment with complements, a cost may exist from the principals not being able to pool their monotonicity constraints. In the cooperative contract regime, (2) indicates that it is possible that one decision function decreases over a range provided that the other increases sufficiently to compensate. Because of the complexity of analyzing the costs of monotonicity constraints on principals under common agency, we do not consider the issue explicitly in this paper, but instead focus attention on environments where the initial cooperative contract is nondecreasing in each argument over Θ .

In order to prove sufficiency in the common agency setting, we will need a modification of assumption A.2 to hold, or alternatively, we can assume A.3 holds for the remainder of this paper. We choose to do the latter.¹³ We are now prepared to provide an equivalent condition for common implementability and feasibility.

Theorem 4 *Any pair of piecewise C^1 decision functions is commonly implementable if and only if $\forall(\hat{\theta}_1, \hat{\theta}_2, \theta) \in \Theta^3$*

$$\begin{aligned} & \int_{\theta}^{\hat{\theta}_2} \int_{\theta}^{\hat{\theta}_1} U_{\hat{\theta}_1, \hat{\theta}_2}(t, s, \theta) dt ds + \int_{\theta}^{\hat{\theta}_2} \int_{\theta}^{\theta} (U_{\hat{\theta}_1, \hat{\theta}_2}(t, s, t) + U_{\hat{\theta}_2, \theta}(t, s, t)) dt ds \\ & + \int_{\theta}^{\hat{\theta}_1} \int_{\theta}^{\theta} (U_{\hat{\theta}_1, \hat{\theta}_2}(s, t, t) + U_{\hat{\theta}_1, \theta}(s, t, t)) dt ds \leq 0, \end{aligned} \quad (1.15)$$

and (1.8) [equivalently, (1.12)] is satisfied. In addition, if and only if (1.11) holds, the contract pair is commonly feasible.

Proof: Following an identical argument to that in the proof of Theorem 2, quasi-linearity guarantees the existence of transfer functions which satisfy (1.12) at all points where $x_i(\theta)$ is differentiable. See Hurewicz [1958, Ch. 2, Theorem 12].

To prove incentive compatibility, we suppose to the contrary that there exists some

¹³Such a modification would require for any $(x_i, t_i, \theta) \in \mathcal{X}^2 \times \mathcal{R} \times \Theta$, there exists a $K > 0$ such that

$$\left\| \left(\frac{U_{x_i}(x_1(\theta), x_2(\theta), t_1 + t_2, \theta)}{U_{t_i}(x_1(\theta), x_2(\theta), t_1 + t_2, \theta)} - \frac{U_{x_i}(x_1(\theta), x_2(\theta), t'_1 + t'_2, \theta)}{U_{t_i}(x_1(\theta), x_2(\theta), t'_1 + t'_2, \theta)} \right) \frac{dx_i(\theta)}{d\hat{\theta}_i} \right\| \leq K \sum_{j=1}^2 \|t_j - t'_j\|,$$

uniformly in x_1, x_2 , and θ , where $\|\varphi\| = \sup_{\theta \in \Theta} |\varphi(\theta)|$.

$(\hat{\theta}_1, \hat{\theta}_2, \theta) \in \Theta^3$ such that $U(\hat{\theta}_1, \hat{\theta}_2, \theta) - U(\theta, \theta, \theta) > 0$. This implies

$$U(\hat{\theta}_1, \hat{\theta}_2, \theta) - U(\hat{\theta}_1, \theta, \theta) + U(\hat{\theta}_1, \theta, \theta) - U(\theta, \theta, \theta) > 0.$$

Integrating we obtain

$$\int_{\theta}^{\hat{\theta}_2} U_{\hat{\theta}_2}(\hat{\theta}_1, s, \theta) ds + \int_{\theta}^{\hat{\theta}_1} U_{\hat{\theta}_1}(s, \theta, \theta) ds > 0.$$

(1.8) implies that $U_{\hat{\theta}_i}(s, s, s) = 0 \forall s \in (\underline{\theta}, \bar{\theta})$, $i = 1, 2$, and so

$$\begin{aligned} & \int_{\theta}^{\hat{\theta}_2} \left[(U_{\hat{\theta}_2}(\hat{\theta}_1, s, \theta) - U_{\hat{\theta}_2}(\theta, s, \theta)) + (U_{\hat{\theta}_2}(\theta, s, \theta) - U_{\hat{\theta}_2}(s, s, s)) \right] ds \\ & + \int_{\theta}^{\hat{\theta}_1} (U_{\hat{\theta}_1}(s, \theta, \theta) - U_{\hat{\theta}_1}(s, s, s)) ds > 0. \end{aligned}$$

Integrating again yields

$$\begin{aligned} & \int_{\theta}^{\hat{\theta}_2} \int_{\theta}^{\hat{\theta}_1} U_{\hat{\theta}_1 \hat{\theta}_2}(t, s, \theta) dt ds + \int_{\theta}^{\hat{\theta}_2} \int_s^{\theta} (U_{\hat{\theta}_1 \hat{\theta}_2}(t, s, t) + U_{\hat{\theta}_2 \theta}(t, s, t)) dt ds \\ & + \int_{\theta}^{\hat{\theta}_1} \int_s^{\theta} (U_{\hat{\theta}_1 \hat{\theta}_2}(s, t, t) + U_{\hat{\theta}_1 \theta}(s, t, t)) dt ds > 0, \end{aligned}$$

which contradicts our initial assumption.

Given (1.8) and A.1, we know the agent's utility is nondecreasing in θ . Together with (1.11) this implies that the participation constraint of the agent is satisfied and the contract pair is feasible. \square

The condition in (1.15) illustrates the additional problems involved in common agency contract design. Under the cooperative contract, providing the contract functions are monotone, the sufficient condition for incentive compatibility is $\mathcal{U}_{x_i \theta} > 0$. This is the Spence-Mirrlees single-crossing property: better types find it marginally cheaper to provide x_i . Under common agency, our first instinct is to suppose that some generalized form of the single-crossing property is sufficient. For example, taking $x_j(\cdot)$ as given, the single-crossing analog in the common agency setting is $\mathcal{U}_{x_i \theta} + \mathcal{U}_{x_1 x_2} x'_j(\theta) > 0$. If

principal i can be assured that principal j 's contract is always incentive compatible (for example, principal j actually observes θ), then this is sufficient, as (1.15) indicates. For instance, take $\hat{\theta}_1 = \theta$ and $\hat{\theta}_2 \neq \theta$. Then only the second term of (1.15) matters, which must be negative if our generalized single-crossing property holds. But even if this general single-crossing property is true for both contracts, the first term in (1.15) may still be positive when $\hat{\theta}_1 \neq \theta \neq \hat{\theta}_2$. In particular, if $\mathcal{U}_{x_1 x_2} < 0$ and $\hat{\theta}_i < \theta < \hat{\theta}_j$, or if $\mathcal{U}_{x_1 x_2} > 0$ and either $\hat{\theta}_1, \hat{\theta}_2 > \theta$ or $\hat{\theta}_1, \hat{\theta}_2 < \theta$, the first term may be sufficiently positive to violate the condition in (1.15).

Unfortunately, unlike the simple monotonicity conditions in the cooperative contracting environment, our global incentive compatibility condition under common agency is complicated. With assumptions restricting the magnitude and sign of various third partial derivatives, however, we can find sufficient conditions for the satisfaction of (1.15). Technically, by restricting the change of $\mathcal{U}_{x_1 x_2}$ when evaluated at different points in the domain of $\Theta \times \mathcal{X}^2$, we can verify (1.15) by using more convenient limits of integration. In our analysis of common agency, the complements case is the simplest to examine as there is an easily discernible set of conditions which are sufficient for the validity of (1.15).

Theorem 5 *Let $\mathcal{U}_{x_1 x_2} > 0$ and $\mathcal{U}_{x_1 x_2 \theta} \leq 0$ for all x_1, x_2, θ . Then any pair of piecewise C^1 contracts for which $x_i'(\theta) \geq 0$ and (1.12) are satisfied is commonly implementable.*

The proof of the theorem is provided in the appendix. Providing that the contracts which we analyze in the complements contracting game have nondecreasing decision functions, the simple condition that $\mathcal{U}_{x_1 x_2}$ does not increase in θ is sufficient for incentive compatibility.

Incentive compatibility with substitutes is more difficult to characterize. Here, we shall also make use of restrictions on $\mathcal{U}_{x_1 x_2}$, but we shall use slightly stronger restrictions to obtain a characterization theorem.

Theorem 6 *Let $\mathcal{U}_{x_1 x_2} < 0$ and suppose the cross-partial derivatives of \mathcal{U} are constant (i.e., $\mathcal{U}_{x_1 x_2}(x_1, x_2, \theta) = u_{12}$, $\mathcal{U}_{x_1 \theta}(x_1, x_2, \theta) = u_{1\theta}$, and $\mathcal{U}_{x_2 \theta}(x_1, x_2, \theta) = u_{2\theta}$). Then the*

necessary conditions in (1.12)-(1.14) are sufficient for common implementability if x_1 and x_2 are nondecreasing.

The proof for this theorem is also provided in the appendix. Note that the above conditions on $\mathcal{U}_{x_1 x_2 \theta}$ in both theorems are not necessary for incentive compatibility and are only used for convenience. To the extent that an agent's utility (e.g., production function, etc.) is satisfactorily approximated by a second-order Taylor expansion, we may rest content with the above simplifications. If not, utility functions with higher order terms may be dealt with by a direct check on the integral conditions contained in (1.15).

1.3.2 Strategic Revelation Effects

We now turn to an examination of the conditions for Nash equilibrium in contracts in the principal's contracting game. We initially note that each principal will typically attempt to induce the agent to report falsely to her rival and thereby extract a larger share of the agent's information rents. In equilibrium, all contracts are incentive compatible so that such attempts are useless, but their possibility imposes constraints on the set of equilibrium contracts.

If instead of studying direct-revelation mechanisms we analyzed nonlinear (tax) schedules, $t_i : \mathcal{X} \mapsto \mathfrak{R}$, the rent-competition effect can be thought of as follows. Principal 1 may decide to change her nonlinear schedule in such a way so as to induce a type- θ agent to choose a contract pair, $\{x'_2, t'_2\}$, from Principal 2 meant for type- θ' , – a choice which Principal 2 had not originally intended. In this manner, Principal 1 may act as an accomplice in helping the agent retain additional information rents from Principal 2. Some of these additional rents are, in turn, extracted by Principal 1's new contract.

If we wish to use the direct revelation mechanism design methodology in the common agency setting, we must introduce additional constructions. Suppose that the decision functions are continuous and U is strictly concave in reports so that we may define the

following functions:¹⁴

$$\hat{\theta}_1[\hat{\theta}_2, \theta | x_1(\cdot), x_2(\cdot), t_1(\cdot), t_2(\cdot)] = \arg \max_{\theta'} U(\theta', \hat{\theta}_2, \theta),$$

$$\hat{\theta}_2[\hat{\theta}_1, \theta | x_1(\cdot), x_2(\cdot), t_1(\cdot), t_2(\cdot)] = \arg \max_{\theta'} U(\hat{\theta}_1, \theta', \theta).$$

Note the functional dependence of each $\hat{\theta}_i$ on the mechanisms offered to the agent. Holding Principal 2's contract fixed, a change in Principal 1's contract will affect the report of the agent to Principal 2. For notational ease, we will at times write $\hat{\theta}_1[\theta | x_2(\theta)]$ and $\hat{\theta}_2[\theta | x_1(\theta)]$, since agent preferences are quasi-linear; with such notation it is understood that the offering principal's contract is incentive compatible (i.e., $\hat{\theta}_2 = \theta$ in the first case, and $\hat{\theta}_1 = \theta$ in the second case). Of course, each function depends on all elements of both contracts even though notationally we have only explicitly recognized dependence on the offering firm's decision function.

In our direct-revelation Nash equilibrium contracting environment, each principal chooses her contract offer taking the offer of her rival as fixed. When maximizing over decision functions, the principal also considers the effect of her contract on the agent's choice from her rival's contract.

In equilibrium all contracts are incentive compatible and we can characterize the effect of a change in one principal's contract on the reports of the agent to the other principal.

Theorem 7 *In any pure-strategy differentiable Nash equilibrium, $\forall \hat{\theta}_1, \hat{\theta}_2 \in (\underline{\theta}, \bar{\theta})$ with x_j strictly increasing*

$$\frac{\partial \hat{\theta}_j[\theta | x_i(\theta)]}{\partial x_i} = \mathcal{U}_{x_1 x_2}(x_1(\theta), x_2(\theta), \theta) / [\mathcal{U}_{x_j \theta}(x_i(\theta), x_j(\theta), \theta) + \mathcal{U}_{x_1 x_2}(x_i(\theta), x_j(\theta), \theta) x'_i(\theta)].$$

If $x'_j(\theta) = 0$, then $\frac{\partial \hat{\theta}_i[\cdot]}{\partial x_i} = 0$.

Proof: Suppose that $\{x_1, x_2, t_1, t_2\}$ is a piecewise C^1 pure-strategy Nash equilibrium. Then we know that the agent's first-order condition (1.12) holds for each principal's

¹⁴The continuity of the decision functions is implied by the strict concavity of each principal's pointwise objective function together with a few technical assumptions, which we take up in the next two sections. For now, however, we take continuity as given.

contract for all but a finite set of θ . Fix firm 1's contract and consider the effect of a change in firm 2's menu. A necessary condition for $\hat{\theta}_1$ to be chosen by the agent given his true type is θ and principal 2 contracts with the type- θ agent for $x_2(\theta)$ is that

$$t'_1(\hat{\theta}_1) = -\mathcal{U}_{x_1}(x_1(\hat{\theta}_1), x_2(\theta), \theta)x'_1(\hat{\theta}_1).$$

From (1.12), this condition becomes

$$\left(\mathcal{U}_{x_1}(x_1(\hat{\theta}_1), x_2(\theta), \theta) - \mathcal{U}_{x_1}(x_1(\hat{\theta}_1), x_2^e(\hat{\theta}_1), \hat{\theta}_1)\right) x'_1(\hat{\theta}_1) = 0,$$

where principal 1 expects principal 2 to offer $x_2^e(\theta)$ in her contract with the agent. If x_1 strictly increases in θ , the bracketed expression must be equal to zero. Totally differentiating this expression with respect to $x_2(\theta)$ and $\hat{\theta}_1$ yields

$$\begin{aligned} \mathcal{U}_{x_1 x_2}(x_1(\hat{\theta}_1), x_2(\theta), \theta) dx_2 = \\ \left[\mathcal{U}_{x_1 x_1}(x_1(\hat{\theta}_1), x_2^e(\hat{\theta}_1), \hat{\theta}_1) x'_1(\hat{\theta}_1) - \mathcal{U}_{x_1 x_1}(x_1(\hat{\theta}_1), x_2(\theta), \theta) x'_1(\hat{\theta}_1) + \right. \\ \left. \mathcal{U}_{x_1 x_2}(x_1(\hat{\theta}_1), x_2^e(\hat{\theta}_1), \hat{\theta}_1) x_2^{e'}(\hat{\theta}_1) + \mathcal{U}_{x_1 \theta}(x_1(\hat{\theta}_1), x_2^e(\hat{\theta}_1), \hat{\theta}_1) \right] d\hat{\theta}_1. \end{aligned}$$

In a pure-strategy Nash equilibrium, $x_2(\theta) = x_2^e(\theta)$ and, without loss of generality, the agent tells the truth to each principal so we evaluate this total differential at $\hat{\theta}_1 = \hat{\theta}_2 = \theta$. Simplification immediately results in the expression of the theorem. When $x_i(\theta)$ is constant, a local change in x_2 can have no effect on $\hat{\theta}_1$, and so $\frac{\partial \hat{\theta}_1[\cdot]}{\partial x_2} = 0$. \square

The expressions in Theorem 7 represent the marginal effect that an increase in one principal's contract menu has on the revelation of the agent to the principal's rival. By characterizing the effects of one contract on the incentive compatibility of another, the expression in Theorem 7 will greatly facilitate our search for Nash equilibria in the contract game. One caveat, however, must be made. The validity of Theorem 7 is restricted to the interior of Θ . As a consequence, each principal must additionally consider whether there is a gain to inducing the agent to choose the corner contract from her rival's offer. In a Nash equilibrium, a principal must not find it beneficial to create

bunching at the corner of her rival's contract, where the agent's first-order conditions may not hold with equality. With complements, this will not be a concern; in the case of substitutes, we will require an additional assumption.

Theorem 7 implies that if (1.15) holds and the decision functions are increasing, then the sign of the report function's derivative is the same as the sign of $\mathcal{U}_{x_1 x_2}$. In an equilibrium with complementary goods, an increase in the contracted activity by one principal will result in an increase in the activity of the other by inducing the revelation of a higher type. The reverse is true when the goods are substitutes. Consequently, examining the cases of complements and substitutes separately is in order.

1.4 Analysis of Equilibria with Contract Complements

By decision complements we mean that $\mathcal{U}_{x_1 x_2} > 0$ for all values of x_1, x_2 and θ . That is, an increase in x_1 raises the marginal value (or lowers the marginal cost) of an increase in x_2 . Situations in which the agent's technology possesses economies of scope or positive spillovers (e.g., learning by doing) are cases where an analysis of contract complements is appropriate. We will need an additional technical requirement before we present a partial characterization of the pure strategy Nash equilibrium contracts set.

Assumption 4'' *The following function is globally strictly concave and has an interior maximum over x_i for $i=1,2, \forall \theta \in \Theta$ and for $x_j(\theta), t_j(\theta)$:*

$$\mathcal{V}^i(x_i) + \mathcal{U}(x_i, x_j(\hat{\theta}_j[\theta|x_i]), \theta) - \frac{1-F(\theta)}{f(\theta)} \frac{\partial \mathcal{U}(x_i, x_j(\hat{\theta}_j[\theta|x_i]), \theta)}{\partial \theta} + t_j(\hat{\theta}_j[\theta|x_i]),$$

where $\frac{\partial \hat{\theta}_i[\theta|x_i]}{\partial x_i} = \frac{\mathcal{U}_{x_1 x_2}}{\mathcal{U}_{x_j \theta} + \mathcal{U}_{x_1 x_2} x_i'}$, and $\frac{\partial^2 \hat{\theta}_j[\theta|x_i]}{\partial x_i^2} = \frac{\partial \hat{\theta}_j[\theta|x_i]}{\partial x_i} \left(\frac{\mathcal{U}_{x_1 x_2 x_i}}{\mathcal{U}_{x_1 x_2}} + \mathcal{U}_{x_1 x_2 x_j} x_j' \right)$. In addition, we restrict $t_i \geq 0$, and assume that $\mathcal{V}^i(x_i) + \mathcal{U}(x_1, x_2, \theta) - \frac{1-F(\theta)}{f(\theta)} \mathcal{U}_\theta(x_1, x_2, \theta) \geq 0$ and $\mathcal{U}_{\theta\theta}(x_1, x_2, \theta) \leq 0$ for all $\theta \in \Theta$ and for all $x_1, x_2 \in \mathcal{X}^2$.

Assumption 4'' is a modification of A.4(a) which guarantees us that each principal's maximization program will be pointwise concave in x_i and involve some positive trade

with the agent. Consequently, our concerns with θ_i^* in Section 2 will not arise. Even a zero contribution (negative transfers are not allowed) by principal i will not result in principal j refusing to serve some types in Θ .¹⁵ A.4'' may have to be checked ex post, as the condition depends upon the signs and magnitudes of third-order partial derivatives and the decision functions' derivative which in turn depend endogenously on the choice of x_i by each principal. The assumption, however, is met whenever the full-information maximization program is sufficiently concave and the degree of uncertainty about θ is small. Alternatively, it is also sufficient if the agent's utility function is quadratic in x_1 and x_2 .

We are now prepared to obtain results for equilibrium existence and characterization.

Proposition 3 *Suppose the contracting activities are complements, A.4'' is satisfied, and $\mathcal{U}_{x_1 x_2 \theta} \leq 0$. Furthermore, suppose a pair of decision functions exists which satisfies the following system of differential equations such that $x_i'(\cdot) \geq 0, \forall x_1, x_2, \theta \in \Theta$:*

$$\mathcal{V}_{x_i}^i(x_i) + \mathcal{U}_{x_i} = \frac{1 - F(\theta)}{f(\theta)} \left[\mathcal{U}_{x_i \theta} + \mathcal{U}_{x_j \theta} x_j'(\theta) \mathcal{U}_{x_1 x_2} \left(\mathcal{U}_{x_j \theta} + \mathcal{U}_{x_1 x_2} x_i'(\theta) \right)^{-1} \right]. \quad (1.16)$$

Given our suppositions, these decision functions constitute a pure-strategy differentiable Nash equilibrium of the common agency contracting game. In such a case, the transfer functions satisfy for $i = 1, 2$,

$$t_i(\theta) = \int_{\underline{\theta}}^{\theta} \frac{\partial \mathcal{U}(x_1(s), x_2(s), s)}{\partial x_i} x_i'(s) ds + \alpha_i \mathcal{U}(x_1(\underline{\theta}), x_2(\underline{\theta}), \underline{\theta}), \quad (1.17)$$

for some α_i such that $\alpha_1 + \alpha_2 = 1$.

The proof is presented in the appendix. Unfortunately, we cannot generally show that a nondecreasing solution to the differential equations will exist. Additionally, when such solutions do exist it is quite possible that multiple equilibria arise – differing in both contract levels and transfers– as in Theorem 8 below. We can, however, indicate simple

¹⁵As in Bernheim-Whinston [1986], we wish to focus on equilibria in which positive activity by the agent occurs. As a consequence of intrinsic agency, a Nash equilibrium always exists in which both principals offer contracts which induce non-participation by the agent. We ignore this equilibria in the analysis which follows.

circumstances in which we will indeed have pure strategy differential equilibria.

Theorem 8 *In a symmetric contracting game where $\mathcal{V}^1 \equiv \mathcal{V}^2$ and $\mathcal{U}(s, t, \theta) \equiv \mathcal{U}(t, s, \theta)$, $\forall s, t$, and where, for all x_1, x_2, θ , $(\mathcal{U}_{x_i x_i \theta} + \mathcal{U}_{x_1 x_2 \theta}) \leq 0$, $\mathcal{U}_{x_1 x_2 \theta} \leq 0$, and $\mathcal{U}_{x_i \theta \theta} \leq 0$, a continuum of symmetric differentiable Nash equilibria exist.*

The proof is in the appendix. The conditions on the third derivatives of \mathcal{U} are sufficient, but not necessary. They merely simplify the analysis in the proof. In the case of symmetric equilibria in symmetric games, there is one equilibrium which is Pareto superior from the principals' viewpoint.¹⁶ It is the contract whose contractual offering for $\underline{\theta}$ is equal to $x_i^{coop}(\underline{\theta})$, the contractual offering to the lowest type under the cooperative solution. This contract introduces the least distortion from the cooperative contracts. As we shall see, this contract is also the one most preferred by the agent.

In general, solving for the differential equations in (1.16) is not always straightforward and may require the use of numerical methods. Nonetheless, we can say something about the properties of equilibria which satisfy the differential equations in (1.16) with two corollaries to Proposition 3.

Corollary 4 *Suppose that A.4'(a) holds. The equilibrium contracts in the common agency game with complementary activities have the property that $\forall \theta$, $x_i(\theta) \leq x_i^{coop}(\theta)$, where $x_i^{coop}(\theta)$ is the contract offered by principal i in the cooperative contracting game.*

Proof: Define $\hat{x}_i^{coop}(x_j(\theta), \theta)$ as the solution to

$$K^i(x_1, x_2, \theta) \equiv \mathcal{V}_{x_i}^i(x_i) + \mathcal{U}_{x_i}(x_1, x_2, \theta) - \frac{1 - F(\theta)}{f(\theta)} \mathcal{U}_{x_i \theta}(x_1, x_2, \theta) = 0,$$

which is uniquely defined given the condition on strict concavity in A.4'(a). Thus $x_i^{coop}(\theta) = \hat{x}_i^{coop}(x_j^{coop}(\theta), \theta)$. With nondecreasing contracts, Theorem 3 implies that the

¹⁶There is actually an infinite number of such equilibria, but all share identical decision functions; they differ only in transfers via the choice of α_i . Such a distinction is of minimal economic interest, and so we loosely refer to this situation as one with a *unique* equilibrium.

right hand side of (1.16) is positive, and so $\hat{x}_i^{coop}(x_j(\theta), \theta) \geq x_i(\theta)$. If contracts are symmetric, we are done. Suppose the contracts are not symmetric and that the Corollary is false. That is,

$$\hat{x}_1^{coop}(x_2(\theta), \theta) \geq x_1(\theta) > x_1^{coop}(\theta).$$

Complementarity implies $\frac{\partial \hat{x}_1^{coop}}{\partial x_2} > 0$, and so it is also the case that

$$\hat{x}_2^{coop}(x_1(\theta), \theta) \geq x_2(\theta) \geq x_2^{coop}(\theta).$$

Because $x_i(\theta) > x_i^{coop}(\theta)$, it must be that $K^i(x_i, x_j^{coop}, \theta) < 0$. By (1.16), we know that $K^i(x_i, x_j, \theta) \geq 0$. Thus, by continuity and the mean-value theorem, there exists an \tilde{x}_j such that $K^i(x_i, \tilde{x}_j, \theta) = 0$, where $\tilde{x}_j \in (x_j^{coop}, x_j]$. Similarly, there exists a $\tilde{x}_i \in (x_i^{coop}, x_i]$ such that $K^j(x_j, \tilde{x}_i, \theta) = 0$. We can as a consequence define continuous mappings, $\phi^i : [x_j^{coop}, x_j] \mapsto (x_i^{coop}, x_i]$, for $i = 1, 2$, and hence by Brouwer's Theorem there exists a fixed point that lies in $(x_1^{coop}, x_1] \times (x_2^{coop}, x_2]$, such that $K^1(\tilde{x}_1, \tilde{x}_2, \theta) = K^2(\tilde{x}_1, \tilde{x}_2, \theta) = 0$. But by A.4', there is a unique fixed point which satisfies the first-order conditions for a cooperative contract, and that fixed point differs from $x^{coop}(\theta)$. \square

Corollary 5 *Both principals and the agent weakly prefer the cooperative contract relative to the outcome in the common agency environment.*

Proof: The fact that the two principals are weakly worse off is a trivial implication of the noncooperative setting. To understand the agent's demise, note that the agent's utility is given by

$$U(\theta) = \int_{\underline{\theta}}^{\theta} \frac{\partial \mathcal{U}(x_1(s), x_2(s), s)}{\partial \theta} ds + U(\underline{\theta}).$$

Because $\mathcal{U}_{x_i, \theta} > 0$, the integrand above must be less under common agency than under the cooperative contract (given our result in Corollary 4). \square

Corollary 4 indicates that the distortions introduced by each principal are greater in the common agency environment. The explanation is straightforward. Equation (1.16) has an additional information rent distortion on the right hand side that is not present in

the cooperative contract of Section 2.2. This term represents the rent effect introduced by competition among principals. First, note that there still is no distortion for the agent with type $\bar{\theta}$. Second, since $U_{x_1, x_2} > 0$, the economic activities of the agent are complements, and the distortion introduced by the principals increases. The dependence of the rent effect on the economic nature of the agent's activities is intuitive: In the case of complements, a principal will decrease its exchange of x_i with the agent to attempt to decrease the agent's exchange of x_j with its rival contractor as this allows principal i to elicit truth telling more cheaply from the agent. Of course, in equilibrium each principal attempts to extract as much rent as possible with the result that the competition for the agent's activity decreases the agent's information rents.

The right hand side of (1.4) in Proposition 1 reflects the effect of x_i on the inframarginal rents which must be paid to all types greater than θ . The right hand side of (1.16) in Proposition 3 also reflects the effect of x_i on the inframarginal rents, but the existence of a strategic complementarity adds to $U_{x_i, \theta}$ and increases the rents which must be paid for an increase of x_i . That is, an increase in x_i directly increases the agent's inframarginal rents through $U_{x_i, \theta}$, but it also indirectly increases rents by raising the choice of x_j by the agent, which in turn raises U_θ still further. Hence, in equilibrium the level of x_i is correspondingly lower than in the cooperative case.

Corollary 5 presents another interesting result under common agency with contract complements – all parties are worse off. Common agency makes information rent reduction by each principal more profitable on the margin, which in turn hurts the agent. The conclusion is analogous to the familiar result with product differentiated duopolists competing in prices: when products are complements, each duopolist charges a price in excess of the monopoly price and consumers are harmed by the presence of competition. In our case, the existence of asymmetric information (together with the possibility of secret contracting) prevents the three parties from eliminating the externalities which they impose upon one another.

The work of Laffont and Tirole [1990] on privatization is related to this point. Their model examines the costs and benefits of government ownership of a firm compared to the regulated environment where both the government and stockholders offer managers

noncooperative contracts. The benefits of regulation over public ownership are better incentives for managers to make investments in the firm because the lack of government ownership is a form of commitment not to appropriate managerial inputs. On the other hand, the effect of common agency is to produce less powerful incentive schemes for cost-reducing effort with greater distortions from efficiency. In Laffont-Tirole, however, only one activity by the agent is contractible and there is conflict between the objectives of the government and the stockholders. To understand the intuition of their results regarding the costs of common agency with a single contractible good, consider the case where $\mathcal{U}_{x_1 x_2}(x_1, x_2, \theta) \rightarrow \infty$ (i.e., $\mathcal{U}(x_1, x_2, \theta)$ is approximated by a Leontief function). In such a case, there is effectively one contractual activity and the right hand side of (1.16) approaches $2\mathcal{U}_{x\theta}$. With a single activity under common agency, the resulting distortion from the first-best full information case increases two-fold in absence of payoff dependencies between the principals' objectives.

A final remark about the relationship between intrinsic and delegated agency is in order. When contracting activities are complements, in equilibrium it will never be the case that the agent prefers to contract with only one principal rather than both. As a consequence, there is no loss in generality in examining the case of intrinsic agency for this class of models. With decision complements, it is never profitable for one principal to offer a contract that induces the agent to deal exclusively with her, leaving her competitor without any trade. With complements, we do not have to consider the constraints which an induced exclusive dealing contract would impose on the equilibrium contracts.

1.5 Analysis of Equilibria with Contract Substitutes

Decisions are substitutes when $\mathcal{U}_{x_1 x_2} < 0$ for all x_i and θ . As was the case with our discussion of complements, we do not directly prove the general existence of equilibrium decision contracts which satisfy (1.15). Rather, we make a weaker proposition regarding the characteristics of such functions when they exist in our simple differentiable setting.

Even this is problematic, however, as our previous use of the first-order approach by principal i when considering the effect of her contract on the agent's report to principal

j is questionable without further assumptions. Specifically, it is arguable that principal i may find it worthwhile to induce an agent in some interval of Θ to always choose the corner contract from principal j 's menu (i.e., report either $\underline{\theta}$ or $\bar{\theta}$ to principal j). If, for example, principal i 's ideal offer of x_i for the agent choosing $\underline{x}_j = x_j(\underline{\theta})$ from principal j ' menu is such that the first-order condition for the agent choosing \underline{x}_j is slack, principal i might prefer to induce corner choices by the agent. Such an offer by the principal is not discovered using the first-order approach in her maximization program because the set of incentive compatible allocations may not be a subset of those satisfying the agent's first-order condition for $\underline{\theta}$ and $\bar{\theta}$. This was not a concern in the case of complements where the first-order condition of the agent always binds in an optimal contract. With substitutes, the following assumption is sufficient to remove the problem.

Assumption 5 For all $x_1, x_2, \tilde{x}_1, \tilde{x}_2, \theta$,

$$\frac{\left(1 - \frac{d}{d\theta} \left(\frac{1-F(\theta)}{f(\theta)}\right)\right) \mathcal{U}_{x_1\theta}(x_1, x_1, \theta) - \frac{1-F(\theta)}{f(\theta)} \mathcal{U}_{x_1\theta\theta}(x_1, x_2, \theta)}{\mathcal{V}_{x_1x_1}^1(x_1) + \mathcal{U}_{x_1x_1}(x_1, x_2, \theta) - \frac{1-F(\theta)}{f(\theta)} \mathcal{U}_{x_1x_1\theta}(x_1, x_2, \theta)} \geq \frac{\mathcal{U}_{x_2\theta}(\tilde{x}_1, x_2, \theta)}{\mathcal{U}_{x_1x_2}(\tilde{x}_1, x_2, \theta)},$$

$$\frac{\left(1 - \frac{d}{d\theta} \left(\frac{1-F(\theta)}{f(\theta)}\right)\right) \mathcal{U}_{x_2\theta}(x_1, x_1, \theta) - \frac{1-F(\theta)}{f(\theta)} \mathcal{U}_{x_2\theta\theta}(x_1, x_2, \theta)}{\mathcal{V}_{x_2x_2}^2(x_2) + \mathcal{U}_{x_2x_2}(x_1, x_2, \theta) - \frac{1-F(\theta)}{f(\theta)} \mathcal{U}_{x_2x_2\theta}(x_1, x_2, \theta)} \geq \frac{\mathcal{U}_{x_1\theta}(x_1, \tilde{x}_2, \theta)}{\mathcal{U}_{x_1x_2}(x_1, \tilde{x}_2, \theta)}.$$

Roughly speaking, A.5 requires that the joint surplus of principal i and the agent is sufficiently concave relative to the substitution term, $\mathcal{U}_{x_1x_2}$, and third-order terms are not sufficiently large in absolute value. Following the analysis in Section 4, we can now state the following Proposition.

Proposition 4 Suppose $\mathcal{U}_{x_1x_2} < 0$, A.4'' and A.5 are satisfied. Furthermore, suppose suppose \mathcal{U} has constant cross-partials, and suppose that a pair of piecewise C^1 decision functions exists which satisfies (1.15) and the following system of differential equations, $\forall x_1, x_2, \theta \in \Theta, i = 1, 2$,

$$\mathcal{V}_{x_i}^i(x_i) + \mathcal{U}_{x_i} = \frac{1-F(\theta)}{f(\theta)} \left[\mathcal{U}_{x_i\theta} + \mathcal{U}_{x_i\theta} x_j'(\theta) \mathcal{U}_{x_1x_2} \left(\mathcal{U}_{x_i\theta} + \mathcal{U}_{x_1x_2} x_i'(\theta) \right)^{-1} \right]. \quad (1.18)$$

Given our suppositions, these decision functions constitute a pure-strategy Nash equilib-

rium in the contracting game. Additionally, the transfer functions satisfy for $i = 1, 2$,

$$t_i(\theta) = \int_{\underline{\theta}}^{\theta} \frac{\partial \mathcal{U}(x_1(s), x_2(s), s)}{\partial x_i} x_i'(s) ds + \alpha_i \mathcal{U}(x_1(\underline{\theta}), x_2(\underline{\theta}), \underline{\theta}), \quad (1.19)$$

for some α_i such that $\alpha_1 + \alpha_2 = 1$.

The proof essentially follows Proposition 3 except that we concern ourselves with the possibility that one principal may desire to induce bunching on the corner of her rival's contract. This problem is taken up in the Appendix.

Equation (1.18) has an additional information rent distortion on the right hand side that is not present in the cooperative contract of Section 2.2, which represents the rent effect introduced by competition among principals. There still is no distortion for the agent with type $\bar{\theta}$, and since the economic activities of the agent are substitutes, the distortion introduced by the principals decreases. A principal will increase her exchange of x_i with the agent to decrease the agent's exchange of x_j with her rival as this allows principal i to elicit truth telling more cheaply from the agent. In equilibrium each principal attempts to extract more rents on the margin with the result that the total sum of the extracted information rents is reduced together with an increase in productive efficiency.

The righthand side of (1.18) in Proposition 4 reflects the effect of x_i on the inframarginal rents. Additionally, the existence of a strategic substitutability affords principal i the opportunity to reduce the rents which must be paid to the agent by decreasing x_j . An increase in x_i directly increases the agent's inframarginal rents through $\mathcal{U}_{x_i, \theta}$, but it also indirectly decreases rents by lowering the choice of x_j by the agent, which in turn lowers \mathcal{U}_{θ} .

When both the preferences and the equilibrium contracts in Proposition 4 are symmetric, the equilibrium common agency contract lies almost everywhere above the cooperative contract due to the informational externalities which each principal imposes upon the other. Each principal prefers to offer a more powerful incentive structure to the agent to reduce the agent's activity with her rival and thereby reduce information rents. In equilibrium, the principals offer sufficiently efficient contracts so that on the margin nothing is gained by reducing a rival's activity with the agent. When contracts

are not symmetric, the nature of the distortions from the efficient level is more difficult to ascertain. Along these lines, we have the following corollary.

Corollary 6 *The commonly implementable contract pair in the pure-strategy contract substitutes equilibrium defined by (1.18), if it exists, must necessarily have*

$$x_i(\theta) \in [x_i^{\text{coop}}(x_j(\theta), \theta), x_i^{\text{eff}}(x_j(\theta), \theta)], \forall \theta,$$

where $x_i^{\text{coop}}(x_j(\theta), \theta)$ is the cooperative contract solution by principal i when facing $x_j(\cdot)$, and where $x_i^{\text{eff}}(x_j(\theta), \theta)$ is the efficient full-information contract solution by principal i given $x_j(\cdot)$. Furthermore, providing that for all values of $x_1, x_2, \bar{x}_1, \bar{x}_2, \theta$ the following conditions hold:

$$\frac{\mathcal{V}_{x_1 x_1}^1(x_1, x_2, \theta) + \mathcal{U}_{x_1 x_1}(x_1, x_2, \theta) - \mathcal{U}_{x_1 x_1 \theta}(x_1, x_2, \theta) \frac{1-F(\theta)}{f(\theta)}}{\mathcal{U}_{x_1 x_2}(x_1, x_2, \theta) - \mathcal{U}_{x_1 x_2 \theta}(x_1, x_2, \theta) \frac{1-F(\theta)}{f(\theta)}} \geq \frac{\mathcal{U}_{x_1 \theta}(x_1, \bar{x}_2, \theta)}{\mathcal{U}_{x_2 \theta}(x_1, \bar{x}_2, \theta)}, \quad (1.20)$$

$$\frac{\mathcal{U}_{x_1 \theta}(\bar{x}_1, x_2, \theta)}{\mathcal{U}_{x_2 \theta}(\bar{x}_1, x_2, \theta)} \geq \frac{\mathcal{U}_{x_1 x_2}(x_1, x_2, \theta) - \mathcal{U}_{x_1 x_2 \theta}(x_1, x_2, \theta) \frac{1-F(\theta)}{f(\theta)}}{\mathcal{V}_{x_2 x_2}^2(x_1, x_2, \theta) + \mathcal{U}_{x_2 x_2}(x_1, x_2, \theta) - \mathcal{U}_{x_2 x_2 \theta}(x_1, x_2, \theta) \frac{1-F(\theta)}{f(\theta)}}, \quad (1.21)$$

then the agent is always weakly better off (and the principals are always weakly worse off) with common agency relative to the cooperative solution.

Proof: The principals are necessarily weakly worse off compared to the cooperative contract. By (1.9) in Theorem 3, we know that

$$\frac{\mathcal{U}_{x_j \theta} \mathcal{U}_{x_1 x_2} x_j'(\theta)}{\mathcal{U}_{x_j \theta} + \mathcal{U}_{x_1 x_2} x_i'(\theta)} \leq 0,$$

which in turn implies that $\mathcal{V}_{x_i}^i(x_i) + \mathcal{U}_{x_i} - \frac{1-F(\theta)}{f(\theta)} \mathcal{U}_{x_i \theta} \leq 0$, and so x_i is chosen above the cooperative levels given x_j : $x_i(\theta) \geq x_i^{\text{coop}}(x_j(\theta), \theta)$. By (1.10) in Theorem 3, we know that the righthand side of (1.18) must be nonnegative for all θ , and so x_i is chosen inefficiently low given the choice of x_j . That is, $x_i(\theta) \leq x_i^{\text{eff}}(x_j(\theta), \theta)$.

The agent's rents from the contracting relationship are given by

$$U(\theta) = \int_{\underline{\theta}}^{\theta} \frac{\partial \mathcal{U}(x_1(s), x_2(s), s)}{\partial \theta} ds + U(\underline{\theta}).$$

Because $U_{x_i, \theta} > 0$, a higher level of contracting activity leads to a larger integrand and hence greater rents for the agent. To see that the agent is at least weakly worse off with the cooperative contract, consider as a reference point in \mathcal{X}^2 the cooperative contract for a given θ : $\{x_1^{coop}(\theta), x_2^{coop}(\theta)\}$. The agent's indifference curve through this point has slope $\frac{dx_2}{dx_1} = -\frac{U_{x_1, \theta}}{U_{x_2, \theta}}$. The functions $x_1^{coop}(x_2, \theta)$ and $x_2^{coop}(x_1, \theta)$ also pass through this point, but with slopes of $-\frac{v_{x_1, x_1}^1 + U_{x_1, x_1} - U_{x_1, x_1, \theta} \frac{1-F}{f}}{U_{x_1, x_2} - U_{x_1, x_2, \theta} \frac{1-F}{f}}$ and $-\frac{U_{x_1, x_2} - U_{x_1, x_2, \theta} \frac{1-F}{f}}{v_{x_2, x_2}^2 + U_{x_2, x_2} - U_{x_2, x_2, \theta} \frac{1-F}{f}}$, respectively. As a consequence of our assumptions, the set of all $\{x_1, x_2\}$ which lie above the curves $x_i^{coop}(x_j, \theta)$ are preferred by the agent compared to the cooperative equilibrium. \square

Although the set of all allocations that are Pareto superior (as judged by both principals and the agent) is a convex set supported by the agent's indifference curve and therefore weakly preferred by the agent, we cannot say that the common agency contracts lie in this convex set. It may be that when preferences and equilibria are not symmetric and the degree of substitution between x_1 and x_2 is high, that an equilibrium exists outside this set. As the degree of substitution approaches zero, however, the common agency contracts must become efficient relative to the cooperative equilibrium.

We still have not proven that common agency equilibria as given by (1.18) actually exist. If, however, we are content with second-order approximations to preferences, and if the underlying uncertainty about type is generated by a process whose hazard rate can be approximated by a linear function, we can make considerably stronger statements about the contracting equilibrium, as we can analytically solve for the contracts given by (1.18). First, we posit the following definition.

Definition 3 *We say that a random process belongs to the class of linear inverse hazard rate distributions (LIHRD) if $f(\theta) = \frac{1}{\gamma}(\bar{\theta} - \underline{\theta})^{-\frac{1}{\gamma}}(\bar{\theta} - \theta)^{\frac{1-\gamma}{\gamma}}$.*

A probability distribution that belongs to such a class has a hazard rate given by $\frac{1-F(\theta)}{f(\theta)} = \gamma(\bar{\theta} - \theta)$. Such a family of probability functions contains the uniform distribution ($\gamma = 1$), as well as arbitrarily close approximations to any exponential distribution.¹⁷

¹⁷An exponential function defined by parameter β over $[0, \infty)$ can be approximated in the linear inverse hazard rate family by choosing $\underline{\theta} = 0$ and letting $\gamma \rightarrow 0$, $\bar{\theta} \rightarrow \infty$ while maintaining $\gamma\bar{\theta} = \beta$. The resulting inverse hazard rate is β .

We can now state our result for quadratic preferences.

Theorem 9 *Suppose that the distribution of θ belongs to the LIHRD class with $\gamma > 0$, the preferences of the principals and the agent are quadratic, and*

$$\frac{V_{x_i x_i}^i + U_{x_i x_i}}{U_{x_1 x_2}} \geq 2(1 + \gamma) \frac{U_{x_i \theta}}{U_{x_j \theta}}, \quad i, j = 1, 2, \quad (1.22)$$

then there exists a unique linear pure-strategy Nash equilibrium in the common agency game $\{x_1^{ca}(\cdot), x_2^{ca}(\cdot)\}$ such that the agent is weakly better off and the principals are weakly worse off than under the optimal cooperative contract.

The result is provided in the Appendix.

1.6 Applications

As we have indicated, many contracting environments are confounded by the presence of common agency. When two or more principals find themselves contracting with the same agent, they generally find themselves worse off because of their failure to cooperate and offer a coordinated set of agency contracts. Understanding the nature of these costs is a requisite first step in our understanding of complex common agency arrangements. We have included here two short analyses of economic problems which involve some form of common agency. The treatments are necessarily incomplete, focusing essentially on the cost aspects of common agency, but they illuminate the broad themes contained in this paper. The first economic problem we address is determining the gains from internalizing transactions to eliminate the costs of common agency in market situations. We examine two manifestations of this concern: the gains to downstream manufacturers from coordination of contracts when dealing with a single input supplier, and the benefits to a firm from using an internal sales force rather than contracting with an independent agency. In a second problem area, we consider the situation of two regulators with imperfectly aligned preferences and ask the welfare question of who gains and who loses from fragmented regulatory authority.

1.6.1 Internal versus External Organization

For the purposes of discussion, we refer to exclusive-agent contracts as *internal* contracting arrangements; in contrast, we say common agency transactions are market-based or *external* arrangements. Internal arrangements are contracts where the parties to the agreement can prevent external forces (such as other principals) from interfering with their relationship. External arrangements are characterized by the absence of such protections. For exposition, we consider joint ventures between firms for the supply of inputs and in-house employees (as opposed to outside agents) as two examples of relationships designed to overcome the externalities of common agency. We recognize that a joint venture is neither necessary nor sufficient for cooperative contracting, and in-house labor is neither necessary nor sufficient for exclusive dealing contracts. Although cooperation could arguably be accomplished through simple contracts between the principals, the existence of additional legal obligations and duties to one another imposed by a joint venture may provide a more effective organization. Similarly, the employee relationship may be a more effective form of internal contracting. Masten (1988), for instance, has emphasized that the legal treatment of employment contracts by courts provides more authority to employers over their employees than a firm could ordinarily have over an independent contractor.¹⁸

Much has been written on the question of when firms prefer such restricted internal relationships (joint ventures, exclusive long-term contracting, internal labor markets, etc.) to unrestricted external transactions. Williamson [1985] indicates gains from internal relationships exist when investment is important but capable of being expropriated in a transient market relationship *and* internal arrangements can prevent such opportunism. On the other hand, Williamson notes that such organizational form is plagued with internal contracting costs, bureaucracies, etc., which result in low-powered incentives, in comparison to the market. Williamson concludes that internalization of market activities

¹⁸We do not wish to make too much of these institutional differences between various alternative organizational forms. If one takes the view that any particular organization is simply a set of "standardized contracts" and is distinguished from other organizations only by the terms of those contracts, the interesting questions focus on the costs and benefits of the various possible contractual terms. Our analysis can be thought of as an examination of the economic costs and benefits of exclusive-agent versus common-agent terms.

occurs when the benefits exceed these costs.

As we shall see below, Williamson's claim that internal contracts are less powerful than market schemes is consistent with common agency under contract substitutes. If effort or productivity of an agent is not observable and the agent's activities are partially substitutable between the two principals' projects, market-based (i.e., external) transactions will be associated with high-powered incentives; exclusive-agent contracts will be associated with low-powered incentives. But here the low-powered internal incentives are not the cost of internal organization, but rather the benefit. Without the influence of another principal's contract, the principal will take advantage of low-powered schemes which are more profitable. It is the presence of excessively powerful market-based schemes that drives the choice to internalize transactions. With complements, we find the implication for the power of schemes is reversed. Market-based transactions are low-powered relative to the internal contracts which would be offered. This affords us a test to determine the importance of a *common agency theory* of internal transactions. A comparison of the schemes from internal and external relationships across firms with varying degrees of economies of scope and scale would be telling in this regard.

Economies of Scope and Contract Complements

Consider a very simple model of a vertical supplier relationship where economies of scope exist in input supply. Two downstream manufacturers, $i = 1, 2$, must each contract for a differentiated input, x_i , which has a constant marginal benefit to manufacture i 's profit of unity. That is, each firm's (principal's) preferences can be summarized as

$$\mathcal{V}^i(x_i) = x_i - t_i,$$

for $i = 1, 2$, where t_i is the payment to the supplier. The supplier's (agent's) preferences exhibit complementarity in production: there are economies of scope available in the production of x_1 and x_2 . For example we suppose

$$\mathcal{U}(x_1, x_2, \theta) = -(\vartheta - \theta)[x_1^2 + x_2^2 - \alpha x_1 x_2],$$

where $\theta \in [\underline{\theta}, \bar{\theta}]$, θ has cumulative distribution function $F(\theta)$, $\vartheta > \bar{\theta}$ and α is a measure for the economies of scope. For concreteness, let $[\underline{\theta}, \bar{\theta}] = [0, 1]$ and $F(\theta) = \theta$ (i.e., θ is uniformly distributed on $[0, 1]$), and let $\alpha = 1$ and $\vartheta = 2$. Then following Propositions 1 and 3 we can derive the optimal contracts under full-information, $x_i^{eff}(\theta)$, incomplete information with cooperation, $x_i^{coop}(\theta)$, and incomplete information with common agency, $x_i^{ca}(\theta)$.

Result 1 *The first-best contract and the cooperative contract are given by the following, respectively:*

$$x_i^{eff}(\theta) = \frac{1}{2 - \theta},$$

$$x_i^{coop}(\theta) = \frac{1}{3 - 2\theta}.$$

The Pareto-dominating common agency equilibrium is defined by the following differential equation

$$\frac{dx_i^{ca}(\theta)}{d\theta} = -\frac{x_i^{ca}(\theta)}{2 - \theta} \left[\frac{1 - (3 - 2\theta)x_i^{ca}(\theta)}{1 - (4 - 3\theta)x_i^{ca}(\theta)} \right]$$

with $x_i^{ca}(\underline{\theta}) = 1/3$.

These contracts are illustrated in Figure 1. Here, common agency introduces more variation in the decision variables, although the quantity/quality spectrum remains unchanged under the Pareto superior equilibrium contract.

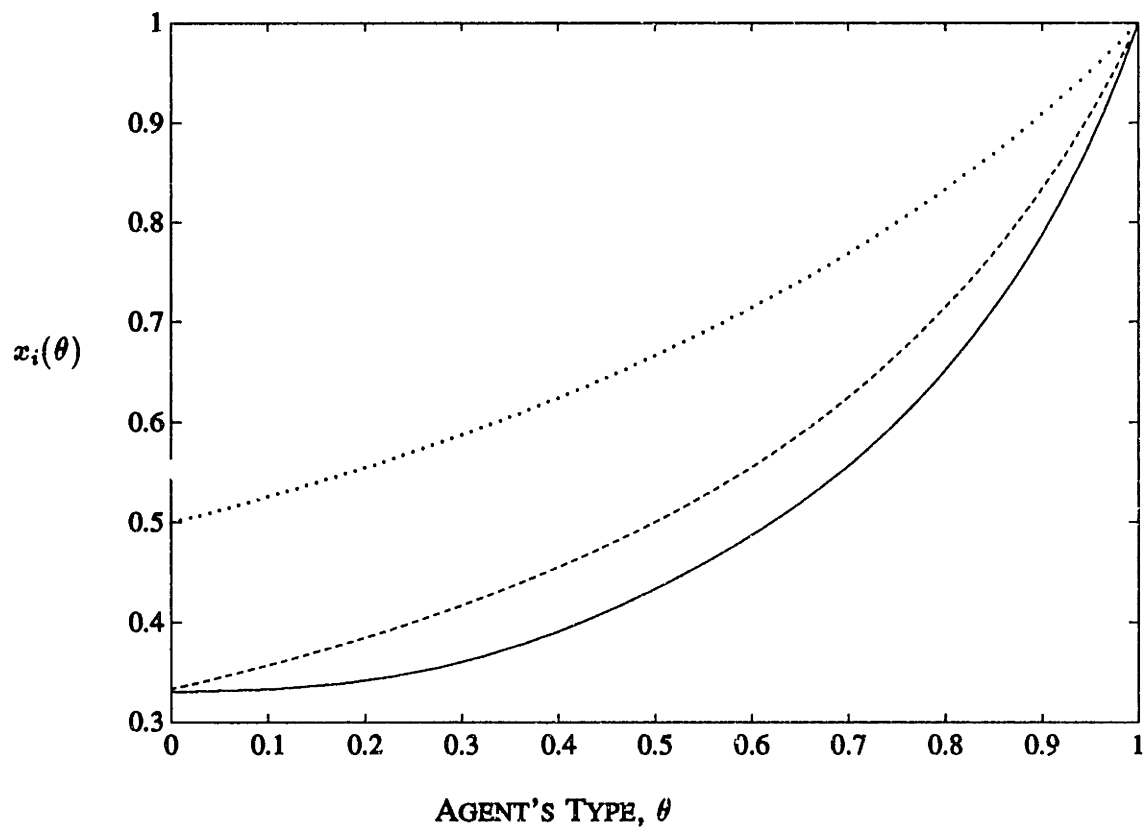
Now consider the decision to internalize the supply transaction. Suppose that principal 1 is already committed to contracting with the agent because of the high cost of alternative arrangements. Principal 2, however, has a choice: she can contract with the same agent, or setup her own input supplier with whom she will exclusively contract. Under the latter internal contracting relationship with an exclusive agent, the agent's preferences are given by

$$U = t_i + \mathcal{U}(x_1, x_2, \theta) = t_i - (\vartheta - \theta)[x_i^2 - \alpha\beta x_1 x_2],$$

where $\beta \in [0, \frac{1}{2}]$ measures the degree to which the principal can capture the scope economies through internal production. If $\beta = \frac{1}{2}$, the principal can convert all of the economies of scope which existed in the common agency framework to economies in

FIGURE 1: CONTRACTUAL RELATIONSHIPS WITH COMPLEMENTARY PRODUCTION

.... Full-Information Contract
---- Cooperative Contract
— Non-Cooperative Contract



producing x_2 alone. Alternatively, one can think of β as the degree to which spillovers continue to occur between two internalized vertical relationships. The unknown marginal cost parameters of each agent are assumed to be independently distributed. We have the following result.

Result 2 *There exists a value of $\beta^* \in (0, \frac{1}{2})$ such that the manager will prefer to internalize production whenever $\beta \geq \beta^*$.*

The result follows from Proposition 3. In the symmetric model under study, $\beta = \frac{1}{2}$ corresponds to the principal obtaining the same level of profits as in the cooperative contracting case. Because there are positive losses associated with common agency in our model, profits under the internalized organization must be greater. Because profits are increasing and continuous in β , we have the result.

Contract Substitutes

Related to this work is that of Holmström and Milgrom [1990] who consider a similar question: When does a manager find it desirable to use an internal sales force rather than an independent contractor to sell her products? Assuming that internal sales forces can be monitored so as to prevent an agent from working for two principals (while an independent agent cannot), they argue that the independent agent's option of exerting effort selling another firm's products may make an internal sales force more desirable.¹⁹ An internal force can be expected to expend a minimal amount of time on the firm's own sales; an independent sales force must be given high-powered incentive schemes to induce the correct level of effort. Their theories regarding the optimal job-task design are closely related to this work on adverse selection. Common agency applied to corporate organization can be thought of as a special case of activity design for an agent; the choice is whether to allow the agent two activities (common agency) or only one (exclusive dealing). With adverse selection and substitutes, the common agency story told here

¹⁹Again, for exposition we have supposed that a firm cannot write an exclusive-dealing contract with an independent agent. Alternatively, we could define agents with exclusive employment contracts to be internal employees and agents with unrestricted contracts to be independent agents without affecting our analysis.

reaches similar conclusions: It may be desirable to exclude the agent from the market in order to allow lower-powered, more profitable contracts. This story, as well as that of Holmström and Milgrom, is consistent with the empirical work by Anderson and Schmittlein [1984]. They find that uncertainty caused by difficulty in equitably measuring individual performance among sales people in the electronics industry is statistically significant in explaining the extent of vertical integration with a firm's sales force.

We now consider a related model which addresses the question: When does a firm find it optimal to use an internal sales force if sales effort is substitutable across the principals' product lines and the productivity of the sales force is private information. Specifically, we consider the case of common agency under adverse selection and moral hazard using a model similar to that in Laffont and Tirole [1986] but with two principals.

Consider a production environment with two risk-neutral firms (principals) and a risk-neutral sales person (agent). The question at issue is the magnitude of the gain that a firm will obtain from internalizing its sales force rather than contracting with a common agent.

The sales force sells units, x_1 and x_2 , for each firm, which are a function of an intrinsic productivity parameter of the agent, θ , and the agent's effort, e_i : $x_i = \theta + e_i$. The agent's cost of effort is convex and quadratic, and efforts are substitutes: $\psi(e_1, e_2) \equiv \frac{1}{2}\psi_{11}e_1^2 + \frac{1}{2}\psi_{22}e_2^2 + \psi_{12}e_1e_2$, where $\psi_{11} > 0$, $\psi_{22} > 0$, and $\psi_{12} > 0$. θ is distributed uniformly on $[0, 1]$.

The payoffs of the two firms are $\mathcal{V}^i(x_i, t_i) \equiv v_i x_i - t_i$, $i = 1, 2$, where t_i is the transfer paid to the employee for the sales of x_i , and v_i is the per unit profit a firm earns on each sale. The firms do not compete on the product market. Their only interactions are through a common agent's marginal costs. Substituting out the agent's effort from his utility function using the sales function results in agent's payoffs that are $U(x_1, x_2, t_1, t_2, \theta) = t_1 + t_2 - \psi(x_1 - \theta, x_2 - \theta)$. With this cost, the full-information efficient contract would set

$$e_i^{eff}(\theta) = x_i^{eff} - \theta_i = \frac{v_i \psi_{jj} - \psi_{12} v_j}{\psi_{11} \psi_{22} - \psi_{12}^2}.$$

In a joint venture, firms can coordinate and offer one contract to the supplier which optimally trades off production distortions against information rent reduction. Following Proposition 1 in Section 2.2, the solution to the joint venture contract is easily derived.

Result 3 *The optimal joint venture contract for a common sales force has $e_i^{coop}(\theta) = x_i^{coop}(\theta) - \theta = e_i^{eff}(\theta) - (1 - \theta)$, $i = 1, 2$, $\forall \theta \in [0, 1]$.*

In order to compare the costs and benefits between using a common agent and using one's own sales force, we need to be precise about the nature of the substitutes under the internal arrangement where one firm uses its own agent exclusively. There are two possible benefits from exclusive agency. First, if firm 1 hired its own exclusive sales force and the agent's cost function remained unchanged, there would be a reduction in sales costs driven by $x_2 = 0$. Second, and more interestingly, more information rents are extracted from the agent absent common agency. In order to focus on the second point, we assume that when a firm employs its own agency, the costs of selling the principal's product are still negatively affected by the level of sales activity undertaken by the other principal's agent. The only change in the environments is that principal i cannot influence agent j 's report to principal j through the choice of her contract. Consequently, the cooperative outcome is identical to the outcome when firms decide to train and employ their own sales force.

When an independent sales force is commonly contracted with by both principals, Theorem 9 provides the following result.²⁰

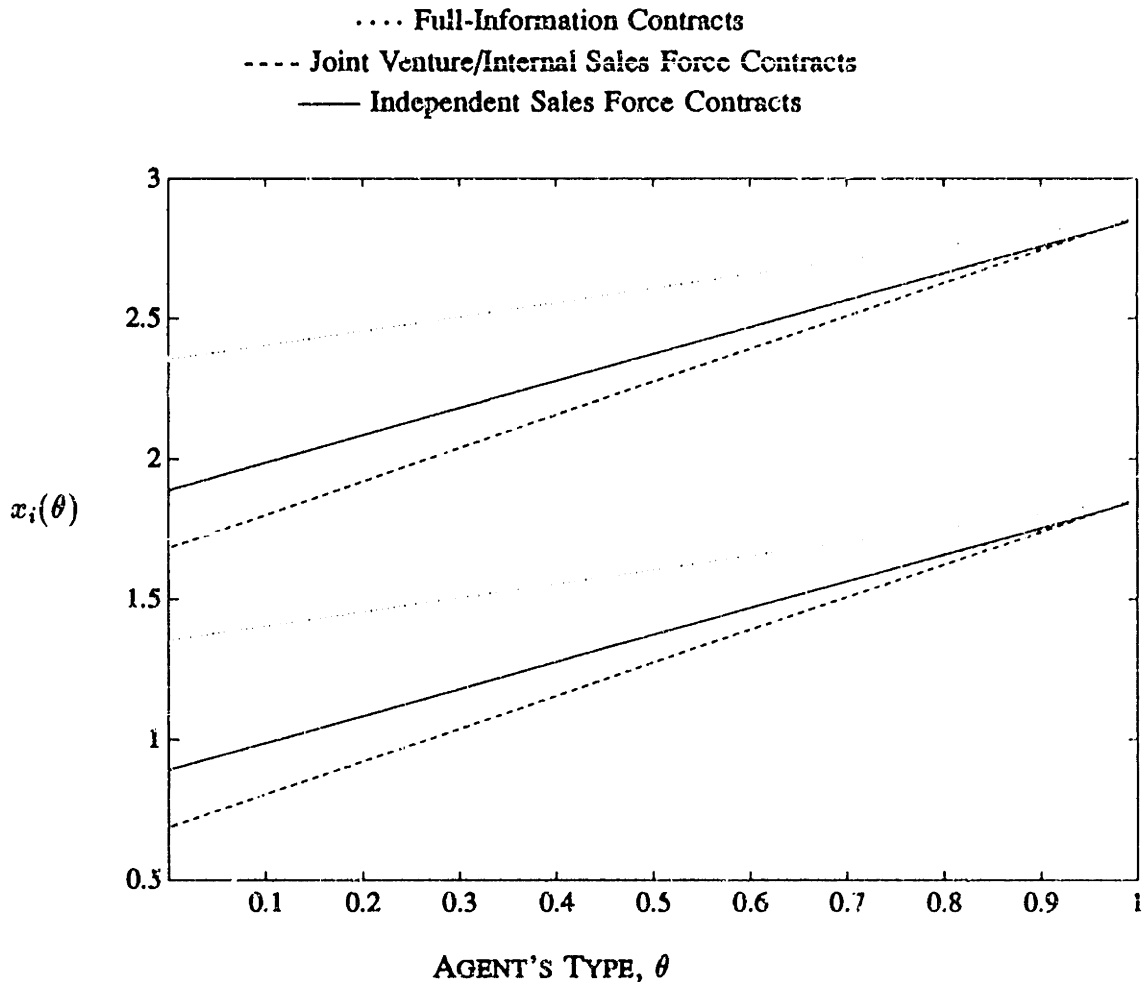
Result 4 *There exists a unique linear pure-strategy Nash equilibrium in the common agency contract game.*

A comparison of the different contracting environments is provided in Figure 2 when parameter values are $\psi_{11} = 3$, $\psi_{22} = 3$, $\psi_{12} = \frac{1}{2}$, $v_1 = 7.5$, and $v_2 = 5$. Firm 1 contracts for a higher level of sales due to its higher per unit profits. Here, because efforts are substitutes, sales are distorted downward more under the internal contracting environment

²⁰We assume that $\psi_{ii}/\psi_{12} \geq 4(\psi_{ii} + \psi_{12})/(\psi_{jj} + \psi_{12})$ for $i, j = 1, 2$, so as to satisfy the conditions of Theorem 9.

than under common agency. This distortion, however, is profit maximizing for the firms. High-powered contracts are less profitable.

FIGURE 2: INTERNAL VERSUS EXTERNAL SALES FORCE.



Now suppose more realistically that there are startup costs, I , to training and employing a sales force. Such costs must be completely born by the firm with an internal sales force, but are shared by both cooperative and noncooperative common-agency principals. Let π_i^{ex} , π_i^{coop} , and π_i^{ca} be the profits, excluding setup costs, associated with an exclusive sales force, joint venture (i.e., cooperative arrangement), and common agent (i.e., noncooperative arrangement), respectively. We know that $\pi_i^{coop} = \pi_i^{ex} > \pi_i^{ca}$. When, however, setup costs are such that $0 < I < (\pi_i^{coop} - \pi_i^{ca})$, the cooperative arrangement is preferred, followed by an exclusive sales force, and then the non-cooperative common agency relationship. As a consequence, when costs are low, even though principals prefer to share

the fixed cost associated with a sales force, they would prefer to expend the extra costs necessary to isolate their agents from their contracting rival when they cannot cooperate.

1.6.2 Multiple Regulators

Consider the problem of two regulatory agencies, each having power to regulate some aspect of an agent's (e.g., a public utility's) operation. This environment is the rule rather than the exception when it comes to administrative law in the United States. Nevertheless, this problem has received little study. One noteworthy exception is the work of Baron [1985]. Baron considers the problem of the dual regulators. In his example, the Environmental Protection Agency (EPA) regulates the level of pollution which a public utility produces and a local Public Utility Commission (PUC) sets rates and production levels for, say, electricity. The EPA has the ability to move first, promulgating some regulation before the PUC has an opportunity to set rates. We consider a simplified version of the same problem, but with simultaneous contracts.

Let x_1 be the level of pollution abatement which the firm achieves. The EPA has simple preferences:

$$v^{EPA}(x_1) = \sqrt{x_1} - t_1.$$

Analogously, the PUC has preferences in accord with local consumers who essentially are unaffected by the utility's pollution (e.g., a coal plant produces acid rain which has no effect on local consumers).

$$v^{PUC}(x_2) = \sqrt{x_2} - t_2.$$

We could, of course, make the preferences of the EPA and the PUC each a function of the firm's profits as well (i.e., make them partially accountable to industry), but we have not done this so as to keep the preferences completely independent.

It is natural to assume that the marginal cost of reducing pollution increases with the level of output. In this case, the contract activities are substitutes. Specifically, let the agent's preferences be like those of the supplier in Section 6.1.2. The agent's

final production of x_i is $e_i + \frac{1}{2}\theta$, where θ is some unknown cost-reducing productivity parameter. We assume that θ is uniformly distributed on $[1, 2]$. The agent's preferences are

$$U = t_1 + t_2 + \mathcal{U}(x_1, x_2, \theta) = t_1 + t_2 - \frac{1}{2}(x_1 + x_2 - \theta)^2.$$

Following Propositions 1 and 4,²¹ we have 2 simple results.

Result 5 *In a cooperative regulatory regime, x_i are chosen to satisfy*

$$\frac{1}{2}x_i^{-\frac{1}{2}} - (x_1 + x_2 - \theta) - (1 - \theta) = 0.$$

Result 6 *In a symmetric equilibrium with fragmented regulation, the EPA and the local PUC choose each x_i in excess of what they would choose with coordinated regulation (i.e., with joint preferences of $\mathcal{V}(x_1, x_2) = \sqrt{x_1} + \sqrt{x_2} - t$). They each choose x_i to satisfy*

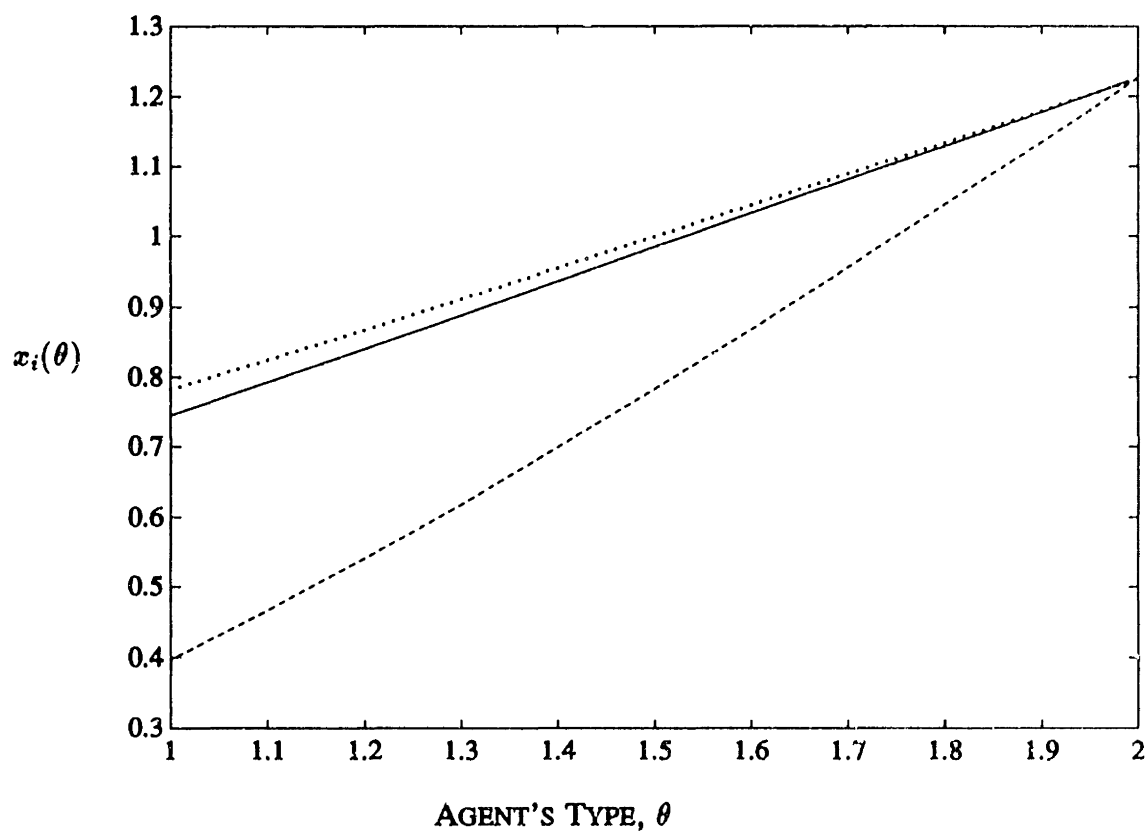
$$\frac{dx_i(\theta)}{d\theta} = -\frac{\frac{1}{2}x_i^{-\frac{1}{2}} - (x_1 + x_2 - \theta) - (1 - \theta)}{\frac{1}{2}x_i^{-\frac{1}{2}} - (x_1 + x_2 - \theta) - 2(1 - \theta)}.$$

This is illustrated in Figure 3. Common agency reduces the distortion in the decision variables which coordinated regulators would otherwise implement. This has several interesting implications for the problem we are examining. First, local consumers and the national constituency for the EPA are worse off. This is due to the costs of common agency. Second, *both the firm and environmentalists are better off* from the high-powered incentive schemes. The firm enjoys more information rents; environmentalists (who we suppose prefer less pollution than the EPA's constituency, perhaps because they pay less taxes) enjoy a more efficient (i.e. lower) level of pollution. This perverse alliance corresponds to that in Laffont-Tirole [1989] where low-powered incentive schemes result from regulatory capture by environmentalists and the regulated firm. In that case both parties gain from collusive arrangements with the regulator to hide information about the firm's costs.

²¹A.5 for Proposition 4 is actually violated in this example. Nonetheless, a numerical examination of the equilibrium contracts reveals that the sufficient condition of $x_i(\underline{x}_j, \theta) \geq x_i^{coop}(\underline{x}_j, \theta)$, for all θ , which is used in the proof to Proposition 4, is met. A.5 is merely a simpler sufficient condition (not necessary) to guarantee that the inequality holds.

FIGURE 3: THE EFFECTS OF MULTIPLE REGULATORS

- Full Information Regulation
- Coordinated Regulation
- Fragmented Regulation



1.7 Conclusion and Further Remarks

Common agency is as prevalent as a lay person's reading of the term would imply. The main focus of this paper has been to develop a theory of techniques to study common agency and to consider the economic effects of common agency on contractual relationships. We have shown that in such environments, if the agent has private information regarding his gains from the contracting activity and the contracting activities in each possible principal-agent relationship are substitutable (complementary), the principals will typically extract less (more) information rents in total and induce less (more) productive inefficiency in the contracting equilibrium than if there were a single principal contracting over the same activities.

The underlying theme of the results presented is that common agency entails costs for the principals. These costs, in turn, can help explain many economic phenomena which we observe. Additionally, as the analysis on substitutes has suggested, common agency may result in high-powered contracts which extract very little of the agent's information rents. Since typical contracting environments have multiple principals, when the contracting activities are substitutes we should expect to see little use of distortionary contracting to reduce information rents. Consequently, even though an environment might be ideal for selection contracts, such contracts may not be observed due to competition. In identical environments with a single principal (e.g., internal organization of a firm), we would expect such contracting schemes. The fact that we do not see many selection contracts may be evidence of healthy competition rather than an oversight by individuals in the marketplace.

Appendix

Proof of Theorem 1: First we show that incentive compatibility implies that $t(\theta)$ is piecewise C^1 . By revealed preference

$$U(\theta + \Delta\theta, \theta + \Delta\theta) - U(\theta + \Delta\theta, \theta) \geq U(\theta + \Delta\theta, \theta + \Delta\theta) - U(\theta, \theta) \geq U(\theta, \theta + \Delta\theta) - U(\theta, \theta).$$

Dividing by $\Delta\theta > 0$ and taking limits as $\Delta\theta \rightarrow 0$ yields $\frac{dU(\theta, \theta)}{d\theta} = U_\theta(x_1, x_2, t, \theta)$. Thus, incentive compatibility implies that the total differential of $U(\theta, \theta)$ exists everywhere. We can use a Taylor expansion at all but a finite number of points and write

$$\begin{aligned} U(\theta + \Delta\theta, \theta + \Delta\theta) - U(\theta, \theta) &= U_{x_1}(x_1, x_2, t, \theta)x'_1(\theta)\Delta\theta + U_{x_2}(x_1, x_2, t, \theta)x'_2(\theta)\Delta\theta \\ &+ U_t(x_1, x_2, t, \theta)\frac{t(\theta + \Delta\theta) - t(\theta)}{\Delta\theta}\Delta\theta + \mathcal{O}(\Delta\theta^2), \end{aligned}$$

for all $\Delta\theta$. Dividing by $\Delta\theta$ we have

$$\frac{t(\theta + \Delta\theta) - t(\theta)}{\Delta\theta} = \frac{U(\theta + \Delta\theta, \theta + \Delta\theta) - U(\theta, \theta)}{\Delta\theta} - U_{x_1}x'_1(\theta) - U_{x_2}x'_2(\theta) - \mathcal{O}(\Delta\theta).$$

The limit on the righthand side exists everywhere but at a finite set of points given the piecewise continuity of $x'_i(\theta)$, and thus $t(\cdot)$ is itself piecewise C^1 .

A necessary condition for maximization by the agent is the satisfaction of first-order and local second-order conditions at $\hat{\theta} = \theta$, at all points of differentiability:

$$U_{\hat{\theta}}(\theta, \theta) = 0,$$

$$U_{\hat{\theta}\hat{\theta}}(\theta, \theta) \leq 0,$$

$\forall \theta \in \Theta$. The first expression immediately gives us (1.1) above. Totally differentiating this expression with respect to θ yields $U_{\hat{\theta}\hat{\theta}}(\theta, \theta) + U_{\hat{\theta}\theta}(\theta, \theta) = 0$, which allows us alternatively to express the local second-order condition as

$$U_{\hat{\theta}\theta}(\theta, \theta) \geq 0,$$

at all points of differentiability. Equivalently,

$$U_{x_1\theta}(x_1, x_2, t, \theta)x'_1(\theta) + U_{x_2\theta}(x_1, x_2, t, \theta)x'_2(\theta) + U_t(x_1, x_2, t, \theta)t'(\theta) \geq 0,$$

for all but a finite set of θ in Θ . Using (1.1) to eliminate $t'(\theta)$ and simplifying yields (1.2). Finally, feasibility implies (1.3) by definition. \square

Proof of Theorem 2: We proceed by showing that there exists a function $t(\cdot)$ satisfying (1.1). Because x_i is piecewise C^1 , there exists a finite set of intervals of Θ on which $(U_{x_i}/U_t)x'_i$ is defined and continuous. Piecewise continuity and the boundedness of x_i allows us to take the closure of these intervals and extend the function over each of these compact subsets of Θ . Following Hurewicz [1958, Ch. 1, Theorem 12], A.2 and $U \in C^2$ implies the existence of a solution which satisfies (1.1) over each of the open intervals, and thus at all points where x_i is differentiable.

To prove that the resulting contracts are globally incentive compatible, suppose otherwise. Let $\hat{\theta} \neq \theta$ be the optimal report for an agent of type θ . By revealed preference, $U(\hat{\theta}, \theta) - U(\theta, \theta) > 0$.

Equivalently,

$$\int_{\underline{\theta}}^{\bar{\theta}} U_{\hat{\theta}}(s, \theta) ds > 0.$$

Using the fact that (1.1) holds everywhere except at a finite number of points yields

$$\int_{\underline{\theta}}^{\bar{\theta}} (U_{\hat{\theta}}(s, \theta) - U_{\hat{\theta}}(s, s)) ds = \int_{\underline{\theta}}^{\bar{\theta}} \int_s^{\theta} U_{\hat{\theta}\theta}(s, t) dt ds > 0.$$

But A.1, together with the assumption of monotonic decision functions, guarantees that $U_{\hat{\theta}\theta}(\hat{\theta}, \theta) \geq 0$, which implies that the preference inequality is violated and we obtain a contradiction. Thus, the contracts are commonly implementable.

Given (1.1) and A.1, the agent's utility is nondecreasing in θ . Therefore, (1.3) is sufficient for participation by the agent and the contracts are feasible. \square

Proof of Proposition 1: Following Mirrlees [1971], we use the agent's indirect utility function: $U(\theta) \equiv \mathcal{U}(x_1(\theta), x_2(\theta), t(\theta))$. Incentive compatibility implies (1.1) which allows us to write

$$U(\theta) = \int_{\underline{\theta}}^{\theta} \mathcal{U}_{\theta}(x_1(s), x_2(s), s) ds + U(\underline{\theta}).$$

A.3(a) implies that $t(\theta) = U(\theta) - \mathcal{U}(x_1(\theta), x_2(\theta), \theta)$, and so A.3(b) implies that the sum of the principals' utilities equals

$$\mathcal{V}^1(x_1(\theta), x_2(\theta)) + \mathcal{V}^2(x_1(\theta), x_2(\theta)) + \mathcal{U}(x_1(\theta), x_2(\theta), \theta) - \int_{\underline{\theta}}^{\theta} \mathcal{U}_{\theta}(x_1(s), x_2(s), s) ds.$$

That is, the principals' joint surplus equals the total gains from trade less information rents which accrue to the agent. Note that partial integration yields

$$\int_{\underline{\theta}}^{\bar{\theta}} \int_{\underline{\theta}}^{\theta} \mathcal{U}_{\theta}(x_1(s), x_2(s), s) f(\theta) ds d\theta = \int_{\underline{\theta}}^{\bar{\theta}} \frac{1 - F(\theta)}{f(\theta)} \mathcal{U}_{\theta}(x_1(\theta), x_2(\theta), \theta) f(\theta) d\theta.$$

From Theorem 2 and A.1, we know that incentive compatibility and agent participation is satisfied if (1.1), (1.3) and monotonicity hold. We have already used (1.1) to substitute out transfers from the maximization problem. Once we obtain the optimal decision functions, we use (1.1) to determine the transfer function, which exists by A.2. This yields (1.5) in the Proposition. It is clear that the maximization of principals' utility requires that (1.3) be binding; it is never profitable to leave information rents to the lowest type agent. Ignoring monotonicity and boundary considerations (i.e., $x_i \in \mathcal{X}$), the principals' relaxed problem reduces to maximizing the expectation of their joint virtual utility

$$\mathcal{V}^1(x_1(\theta), x_2(\theta)) + \mathcal{V}^2(x_1(\theta), x_2(\theta)) + \mathcal{U}(x_1(\theta), x_2(\theta), \theta) - \frac{1 - F(\theta)}{f(\theta)} \mathcal{U}_{\theta}(x_1(\theta), x_2(\theta), \theta).$$

Because the integrand is continuous over a compact set \mathcal{X} , a solution exists for each θ . Maximizing the integrand pointwise in θ yields (1.4) $\forall \theta \in [\theta_i^*, \bar{\theta}]$, and $x_i(\theta) = 0 \forall \theta \in [\underline{\theta}, \theta_i^*]$. A.4 implies that the integrand is globally concave in x_i , and so the first-order conditions are sufficient. Note that the joint virtual utility evaluated at (1.4) is increasing in θ because $\mathcal{U}_{\theta\theta} \leq 0$ (A.4).

Suppose $\theta_1^* \leq \theta_2^*$. Then θ_2^* is chosen to satisfy

$$\mathcal{V}^2(x_1, x_2) + \mathcal{U}(x_1, x_2, \theta_2^*) - \frac{1 - F(\theta_2^*)}{f(\theta_2^*)} \mathcal{U}_\theta(x_1, x_2, \theta_2^*) = 0.$$

This completely defines x_2 over Θ . We now choose θ_1^* to satisfy

$$\mathcal{V}^1(x_1, 0) + \mathcal{U}(x_1, 0, \theta_1^*) - \frac{1 - F(\theta_1^*)}{f(\theta_1^*)} \mathcal{U}_\theta(x_1, 0, \theta_1^*) = 0.$$

This completely determines x_1 over Θ . A similar exercise is used when $\theta_1^* > \theta_2^*$. In either case, the choice of θ_i^* satisfy the conditions of the Proposition.

We now check that the monotonicity and boundary conditions are satisfied. Totally differentiating (1.4), together with A.4, imply that each x_i is nondecreasing in θ , thereby satisfying (1.2). Because each x_i is nondecreasing in θ , A.4(b) implies that the maximum value of each x_i is in \mathcal{X} . \square

Proof of Proposition 2: (Sketch) Proposition 2 follows from the arguments used in Proposition 1, with the exception of proving monotonicity of x_i ; the existence of a single pair of cutoff types, θ_i^* . Supposing that x_1, x_2 satisfy (1.6) over $[\theta_i^*, \bar{\theta}]$, we need to show that $x_i'(\theta) \geq 0$ and that a cutoff point, θ_i^* , exists such that if and only if $\theta \geq \theta_i^*$ are principal i 's profits nonnegative.

(1.6) provides a system of two equations that define x_1, x_2 . Totally differentiating this system with respect to x_1, x_2 , and θ , and using Cramer's rule to solve for $x_i'(\theta)$ yields $x_i'(\theta) \geq 0$ in light of A.4'(b). Furthermore, given the condition in A.4'(c) which requires

$$\left(\mathcal{V}_{x_j}^i - \frac{1 - F}{f}\right) \mathcal{U}_{\theta x_2} x_j'(\theta) - \frac{1 - F}{f} \mathcal{U}_{\theta\theta} + \left[1 - \frac{d}{d\theta} \left(\frac{1 - F}{f}\right)\right] \mathcal{U}_\theta \geq 0,$$

principal i 's objective function increases in θ . \square

Proof of Theorem 5: Following Theorem 4, it is sufficient to show that (1.15) is satisfied for any pair of nondecreasing decision functions. That is,

$$\begin{aligned} \Lambda(\hat{\theta}_1, \hat{\theta}_2, \theta) &\equiv \int_{\theta}^{\hat{\theta}_2} \int_{\theta}^{\hat{\theta}_1} \mathcal{U}_{x_1 x_2}(s, t, \theta) x_1'(s) x_2'(t) ds dt \\ &+ \int_{\theta}^{\hat{\theta}_1} \int_s^{\theta} [\mathcal{U}_{x_1 x_2}(s, t, t) x_1'(s) x_2'(t) + \mathcal{U}_{x_1 \theta}(s, t, t) x_1'(s)] dt ds \\ &+ \int_{\theta}^{\hat{\theta}_2} \int_t^{\theta} [\mathcal{U}_{x_1 x_2}(s, t, s) x_1'(s) x_2'(t) + \mathcal{U}_{x_2 \theta}(s, t, s) x_2'(t)] ds dt \leq 0, \end{aligned}$$

$\forall (\hat{\theta}_1, \hat{\theta}_2, \theta) \in \Theta^3$. Note that we can decompose the first double integral:

$$\begin{aligned} \int_{\theta}^{\hat{\theta}_2} \int_{\theta}^{\hat{\theta}_1} \mathcal{U}_{x_1 x_2}(s, t, \theta) x_1'(s) x_2'(t) ds dt &= \int_{\theta}^{\hat{\theta}_1} \int_{\theta}^{\gamma + \beta s} \mathcal{U}_{x_1 x_2}(s, t, \theta) x_1'(s) x_2'(t) dt ds \\ &+ \int_{\theta}^{\hat{\theta}_2} \int_{\theta}^{\frac{t-\gamma}{\beta}} \mathcal{U}_{x_1 x_2}(s, t, \theta) x_1'(s) x_2'(t) ds dt, \end{aligned}$$

where $\beta = \frac{\hat{\theta}_2 - \theta}{\hat{\theta}_1 - \theta}$, and $\gamma = \theta(1 - \beta)$. Thus,

$$\begin{aligned}\Lambda(\hat{\theta}_1, \hat{\theta}_2, \theta) &= \int_{\theta}^{\hat{\theta}_1} \int_s^{\theta} [\mathcal{U}_{x_1 x_2}(s, t, t)x_2'(t) - \mathcal{U}_{x_1 x_2}(s, t, \theta)x_2'(t) + \mathcal{U}_{x_1 \theta}(s, t, t)] x_1'(s) dt ds \\ &+ \int_{\theta}^{\hat{\theta}_2} \int_t^{\theta} [\mathcal{U}_{x_1 x_2}(s, t, s)x_1'(s) - \mathcal{U}_{x_1 x_2}(s, t, \theta)x_1'(s) + \mathcal{U}_{x_2 \theta}(s, t, s)] x_2'(t) ds dt \\ &+ \int_{\theta}^{\hat{\theta}_1} \int_s^{\gamma + \beta s} \mathcal{U}_{x_1 x_2}(s, t, \theta)x_1'(s)x_2'(t) dt ds + \int_{\theta}^{\hat{\theta}_2} \int_t^{\frac{t-\gamma}{\beta}} \mathcal{U}_{x_1 x_2}(s, t, \theta)x_1'(s)x_2'(t) ds dt.\end{aligned}$$

Integrating yields

$$\begin{aligned}\Lambda(\hat{\theta}_1, \hat{\theta}_2, \theta) &= \int_{\theta}^{\hat{\theta}_1} \int_s^{\theta} \left\{ \mathcal{U}_{x_1 \theta}(s, t, t) + \int_{\theta}^t \mathcal{U}_{x_1 x_2 \theta}(s, t, u)x_2'(t) du \right\} x_1'(s) dt ds \\ &+ \int_{\theta}^{\hat{\theta}_2} \int_t^{\theta} \left\{ \mathcal{U}_{x_2 \theta}(s, t, s) + \int_{\theta}^s \mathcal{U}_{x_1 x_2 \theta}(s, t, u)x_1'(s) du \right\} x_2'(t) ds dt \\ &+ \int_{\theta}^{\hat{\theta}_1} \int_s^{\gamma + \beta s} \mathcal{U}_{x_1 x_2}(s, t, \theta)x_1'(s)x_2'(t) dt ds + \int_{\theta}^{\hat{\theta}_2} \int_t^{\frac{t-\gamma}{\beta}} \mathcal{U}_{x_1 x_2}(s, t, \theta)x_1'(s)x_2'(t) ds dt.\end{aligned}$$

But note that we can combine the last two terms to obtain

$$\begin{aligned}\Lambda(\hat{\theta}_1, \hat{\theta}_2, \theta) &= \int_{\theta}^{\hat{\theta}_1} \int_s^{\theta} \left\{ \mathcal{U}_{x_1 \theta}(s, t, t) + \int_{\theta}^t \mathcal{U}_{x_1 x_2 \theta}(s, t, u)x_2'(t) du \right\} x_1'(s) dt ds \\ &+ \int_{\theta}^{\hat{\theta}_2} \int_t^{\theta} \left\{ \mathcal{U}_{x_2 \theta}(s, t, s) + \int_{\theta}^s \mathcal{U}_{x_1 x_2 \theta}(s, t, u)x_1'(s) du \right\} x_2'(t) ds dt \\ &+ \int_{\hat{\theta}_1}^{\hat{\theta}_2} \int_s^{\hat{\theta}_1} \mathcal{U}_{x_1 x_2}(s, t, \theta)x_1'(s)x_2'(t) dt ds.\end{aligned}$$

Given our assumptions about monotonicity and $\mathcal{U}_{x_1 x_2 \theta}$, it is straightforward to verify that each of the three terms in Λ are necessarily nonpositive. Thus (1.15) is satisfied and the pair of contracts is commonly implementable. \square

Proof of Theorem 6: Following Theorem 4, it is sufficient to show that (1.15) is satisfied under the conditions on \mathcal{U} providing that the necessary conditions in (1.13)-(1.14) are satisfied and each x_i is nondecreasing. That is, we take as given for all θ

$$u_{1\theta} + u_{12}x_2'(\theta) \geq 0,$$

$$u_{2\theta} + u_{12}x_1'(\theta) \geq 0,$$

$$u_{1\theta}u_{2\theta} + u_{12}(u_{1\theta}x_1'(\theta) + u_{2\theta}x_2'(\theta)) \geq 0.$$

First, note that (1.15) can be simplified under our conditions on monotonicity and \mathcal{U} :

$$\begin{aligned}\int_{\theta}^{\hat{\theta}_2} \int_{\theta}^{\hat{\theta}_1} u_{12}x_1'(s)x_2'(t) dt ds + \int_{\theta}^{\hat{\theta}_1} \int_s^{\theta} [u_{12}x_1'(s)x_2'(t) + u_{1\theta}x_1'(s)] dt ds \\ + \int_{\theta}^{\hat{\theta}_2} \int_t^{\theta} [u_{12}x_1'(s)x_2'(t) + u_{2\theta}x_2'(t)] ds dt \leq 0.\end{aligned}$$

Using (1.14),

$$u_{1\theta}x_1'(s) + u_{12}x_1'(s)x_2'(t) \geq -u_{12}\frac{u_{1\theta}}{u_{2\theta}}x_1'(t)x_1'(s),$$

and

$$u_{2\theta}x_2'(t) + u_{12}x_1'(s)x_2'(t) \geq -u_{12}\frac{u_{2\theta}}{u_{1\theta}}x_2'(t)x_2'(s).$$

Using the fact that $u_{12} < 0$, it is sufficient for (1.15) that

$$\int_{\theta}^{\hat{\theta}_2} \int_{\theta}^{\hat{\theta}_1} x_1'(s)x_2'(t)dtds - \int_{\theta}^{\hat{\theta}_1} \int_{s}^{\theta} \frac{u_{1\theta}}{u_{2\theta}}x_1'(t)x_1'(s)dtds - \int_{\theta}^{\hat{\theta}_2} \int_t^{\theta} \frac{u_{2\theta}}{u_{1\theta}}x_2'(t)x_2'(s)dsdt \geq 0.$$

Consider the three terms independently. After simplification,

$$\begin{aligned} \int_{\theta}^{\hat{\theta}_2} \int_{\theta}^{\hat{\theta}_1} x_1'(s)x_2'(t)dtds &= [x_1(\hat{\theta}_1) - x_1(\theta)][x_2(\hat{\theta}_2) - x_2(\theta)], \\ \int_{\theta}^{\hat{\theta}_1} \int_s^{\theta} -\frac{u_{1\theta}}{u_{2\theta}}x_1'(t)x_1'(s)dtds &= -\frac{u_{1\theta}}{2u_{2\theta}}[x_1(\hat{\theta}_1) - x_1(\theta)]^2, \\ \int_{\theta}^{\hat{\theta}_2} \int_t^{\theta} -\frac{u_{2\theta}}{u_{1\theta}}x_2'(t)x_2'(s)dsdt &= -\frac{u_{2\theta}}{2u_{1\theta}}[x_2(\hat{\theta}_2) - x_2(\theta)]^2. \end{aligned}$$

The sum of these expressions forms a binomial which can be simplified to yield

$$\left\{ u_{1\theta}[x_1(\hat{\theta}_1) - x_1(\theta)] - u_{2\theta}[x_2(\hat{\theta}_2) - x_2(\theta)] \right\}^2 \geq 0.$$

Thus, (1.15) holds. □

Proof of Proposition 3: Again we use the agent's indirect utility function:

$$U(\theta) \equiv \mathcal{U}(x_1(\theta), x_2(\theta), \theta) + t_1(\theta) + t_2(\theta).$$

Incentive compatibility implies (1.12), which allows us to write

$$U(\theta) = \int_{\underline{\theta}}^{\theta} \frac{\partial \mathcal{U}(x_1(s), x_2(s), s)}{\partial \theta} ds + U(\underline{\theta}).$$

A.3 implies that $t_1(\theta) + t_2(\theta) = U(\theta) - \mathcal{U}(x_1(\theta), x_2(\theta), \theta)$. We first analyze the problem of Principal 1. A.3(b) implies that her gain from an incentive compatible exchange (but not necessarily an exchange which is incentive compatible for her rival's contract) is

$$\begin{aligned} \mathcal{V}^1(x_1(\theta)) + \mathcal{U}(x_1(\theta), x_2(\hat{\theta}_2[\theta|x_1(\theta)]), \theta) - \int_{\underline{\theta}}^{\theta} \frac{\partial \mathcal{U}(x_1(s), x_2(\hat{\theta}_2[s|x_1(s)]), s)}{\partial \theta} ds \\ + t_2(\hat{\theta}_2[\theta|x_1(\theta)]) - U(\underline{\theta}), \end{aligned}$$

providing that $\hat{\theta}_2 \in (\underline{\theta}, \bar{\theta})$.

Principal 1's surplus equals the total gains from trade in x_1 less information rents which accrue to the agent plus the agent's compensation from principal 2. Partial integration allows us

to conclude

$$\int_{\underline{\theta}}^{\bar{\theta}} \int_{\underline{\theta}}^{\theta} \frac{\partial \mathcal{U}(x_1(s), x_2(\hat{\theta}_2[s|x_1(s)]), s)}{\partial \theta} f(\theta) ds d\theta = \int_{\underline{\theta}}^{\bar{\theta}} \frac{1-F(\theta)}{f(\theta)} \frac{\partial \mathcal{U}(x_1(\theta), x_2(\hat{\theta}_2[\theta|x_1(\theta)]), \theta)}{\partial \theta} f(\theta) d\theta.$$

From Theorem 5 we know that incentive compatibility and agent participation are satisfied provided that (1.12), and monotonicity hold. We have already used (1.12) to substitute out transfers from the maximization problem. Once we obtain the optimal decision functions, we use (1.12) to determine the transfer function.

$$\int_{\underline{\theta}}^{\bar{\theta}} \left\{ \mathcal{V}^1(x_1(\theta)) + \mathcal{U}(x_1(\theta), x_2(\hat{\theta}_2[\theta|x_1(\theta)]), \theta) - \frac{1-F(\theta)}{f(\theta)} \frac{\partial \mathcal{U}(x_1(\theta), x_2(\hat{\theta}_2[\theta|x_1(\theta)]), \theta)}{\partial \theta} + t_2(\hat{\theta}_2[\theta|x_1(\theta)]) - U(\underline{\theta}) \right\} f(\theta) d\theta.$$

By A.4'', the solution to the relaxed program can be found by differentiating the integrand pointwise in θ and setting the result equal to zero yielding (1.16) if we can be certain that principal 1 finds it optimal that $\hat{\theta}_2 \in (\underline{\theta}, \bar{\theta})$, for all $\theta \in (\underline{\theta}, \bar{\theta})$.

To see that bunching at $\underline{\theta}$ is not optimal for principal 1, consider the functions $x_1^{coop}(\underline{x}_2, \theta)$ and $\check{x}_1(\underline{x}_2, \theta)$. The first function is defined as the value of x_1 which principal 1 would prefer to choose if principal 2 always offered $\underline{x}_2 = x_2(\underline{\theta})$. The second is the maximum value of x_1 which principal 1 can offer to agent θ in order to induce the agent to choose the $\{x_1(\theta), \underline{x}_2\}$ allocation. If $x_1^{coop}(\underline{x}_2, \theta) \geq \check{x}_1(\underline{x}_2, \theta)$, then the constraint facing the principal who wishes to induce bunching at $\underline{\theta}$ must be binding. If it binds, the first-order condition of the agent is satisfied, and the program above which uses the first-order approach is valid. To see that the sufficient inequality above holds, note that at $\underline{\theta}$ it must be the case (since $\mathcal{U}_{x_1 x_2} > 0$) that $x_1^{coop}(\underline{x}_2, \underline{\theta}) \geq \check{x}_1(\underline{x}_2, \underline{\theta})$. Furthermore, under our assumptions in A.4'', x_1^{coop} is increasing in θ while \check{x}_1 is decreasing in θ . Thus, the desired inequality holds. A similar argument establishes that with complements, bunching is never optimal at $\bar{\theta}$.

Given that a nondecreasing solution to (1.16) exists, our assumption that $\mathcal{U}_{x_1 x_2 \theta} \leq 0$ implies that the contracts are commonly implementable. A.4'' implies that an α_i exists such that it is optimal for all θ to be served by each principal. Providing that the transfers are chosen as in the Proposition, the contracts are globally incentive compatible and individually rational. \square

Proof of Theorem 8: The proof follows directly from Proposition 3, except that we must additionally show that a continuum of symmetric, nondecreasing solutions to the differential equations in (1.16) exists. Define $s(x, \theta) \equiv \mathcal{V}_x(x) + \mathcal{U}_x(x, x, \theta)$ and define the surface

$$\mathcal{D} \equiv \{x, \theta | \theta \in [\underline{\theta}, \bar{\theta}), N(x, \theta) \geq 0, D(x, \theta) < 0\},$$

where $N(x, \theta) \equiv s(x, \theta) - \frac{1-F(\theta)}{f(\theta)} \mathcal{U}_\theta(x, x, \theta)$ and $D(x, \theta) \equiv s(x, \theta) - 2 \frac{1-F(\theta)}{f(\theta)} \mathcal{U}_\theta(x, x, \theta)$. Our assumptions on \mathcal{U} imply that there is a unique x for each θ such that $N(x, \theta) = 0$; this point lies in \mathcal{D} , and so the latter is nonempty. Furthermore, our assumptions imply that $\frac{\partial D}{\partial x} < 0$, and that along the curve defined by $N(x, \theta) = 0$ we have $\frac{dx}{d\theta} > 0$. As a consequence, we have the curve given by N lying above the curve given by $D(x, \theta) = 0$ over the domain of Θ , with the former having positive slope everywhere.

Manipulating the differential equation given in Proposition 3 and using symmetry implies that

$$x'(\theta) = -\frac{U_{x\theta}(x, x, \theta) N(x, \theta)}{U_{x_1 x_2}(x, x, \theta) D(x, \theta)}.$$

Thus, if a differential equation exists in \mathcal{D} , it necessarily has the desired monotonicity property. Choose any point in \mathcal{D} and consider its direction of movement. It cannot cross the $N(x, \theta)$ locus from below, as the derivative in the neighborhood of N is 0 and the locus N has strictly positive slope. It cannot cross the D locus from above as $x' \rightarrow +\infty$ as x approaches D and the locus of points satisfying D has finite slope. Thus, any point in \mathcal{D} remains in \mathcal{D} ; and moreover, in any neighborhood, x' locally satisfies a Lipschitz condition. Following Hurewicz [1958, Chapter 2, Theorem 12], a global differential equation exists which satisfies the equation in Proposition 3. Additionally, such an equation exists for any initial point in the half-open interval $\mathcal{D}(\underline{\theta}) = \{x, \underline{\theta} | N(x, \underline{\theta}) \geq 0, D(x, \underline{\theta}) < 0\}$. We thus have a continuum of nondecreasing solutions. \square

Proof (Sketch) of Proposition 4: Proposition 4 follows from the analysis of Proposition 3, except in so far as we must check that A.5 is sufficient for corner bunching to be suboptimal.

First, note that bunching will never occur at $\bar{\theta}$. If principal 1 chooses to induce bunching by the agent on Principal 2's contract, the higher level of induced x_2 will result in both more information rents being paid to the agent by Principal 1, as well as reduced profits from lower purchases from the agent.

Second, consider bunching at $\underline{\theta}$. As in the proof to Proposition 3, it is sufficient to show that $x_1^{coop}(\underline{x}_2, \theta) \leq \check{x}_1(\underline{x}_2, \theta)$, where $\check{x}_1(\underline{x}_2, \theta)$ is now the *minimum* value of x_1 which principal 1 can offer to agent θ in order to induce the agent to choose the $\{x_1(\theta), \underline{x}_2\}$ allocation. In such a case the constraint facing the principal who wishes to induce bunching at $\underline{\theta}$ must be binding. If it binds, the first-order condition of the agent is satisfied, and the program above which uses the first-order approach is valid. To see that the sufficient inequality above holds, note that at $\underline{\theta}$ it must be that (since $U_{x_1 x_2} < 0$) $x_1^{coop}(\underline{x}_2, \underline{\theta}) < \check{x}_1(\underline{x}_2, \underline{\theta})$. Furthermore, under our assumptions in A.5, x_1^{coop} is increasing at a slower rate in θ than is \check{x}_1 . Thus, the desired inequality holds. \square

Proof of Theorem 9: Define the quadratic preferences as follows:

$$V^i(x_i) \equiv v_0^i + v_i x_i + \frac{1}{2} v_{ii} x_i^2, i = 1, 2,$$

$$U(x_1, x_2, \theta) \equiv u_0 + (u_1 + u_{1\theta}\theta)x_1 + (u_2 + u_{2\theta}\theta)x_2 + u_{12}x_1x_2 + \frac{u_{11}}{2}x_1^2 + \frac{u_{22}}{2}x_2^2.$$

We look for linear solutions of the form $x_i = \bar{x}_i^{eff} - \lambda_i(\underline{\theta} - \theta)$, where \bar{x}_i^{eff} is the efficient allocation given that $\theta = \bar{\theta}$. From Theorem 6 and Proposition 4, we need only show that of the linear solutions to (1.18), there is a unique pair $\{\lambda_1, \lambda_2\}$ such that each $\lambda_i \geq 0$ and (1.13)-(1.14) are satisfied (i.e., $\lambda_i \leq -\frac{u_{i\theta}}{u_{12}}$ and $u_{1\theta}u_{2\theta} + u_{12}(u_{1\theta}\lambda_1 + u_{2\theta}\lambda_2) \geq 0$).

Substituting the candidate linear solutions into (1.18) and simplifying yields, for $i = 1, 2$,

$$\begin{aligned} v_i + v_{ii}(\bar{x}_i^{eff} - \lambda_i(\bar{\theta} - \theta)) + u_i + u_{i\theta}\theta + u_{12}(\bar{x}_j^{eff} - \lambda_j(\bar{\theta} - \theta)) + u_{ii}(\bar{x}_i^{eff} - \lambda_i(\bar{\theta} - \theta)) \\ = \gamma(\bar{\theta} - \theta) \left(u_{i\theta} + \frac{u_{j\theta}\lambda_j u_{12}}{u_{j\theta} + u_{12}\lambda_j} \right). \end{aligned}$$

This expression must hold for any θ .

First, note that if some $\lambda_i = 0$, the above expression cannot be true. For example, $\lambda_1 = 0$ implies that $\lambda_2 = -\frac{u_{1\theta}}{u_{12}}$, but then the optimal choice for $\lambda_1 \neq 0$. Hence, no fixed point can contain a zero component and we can treat λ_i as a nonzero number. Second, note that since the above expression must hold true for all θ , a necessary and sufficient condition for $\{\lambda_1, \lambda_2\}$ is that the coefficients of θ sum to zero. That is, for $i = 1, 2$,

$$-(u_{j\theta} + u_{12}\lambda_i)(u_{i\theta} + u_{12}\lambda_j + (v_{ii} + u_{ii})\lambda_i) = \gamma(u_{1\theta}u_{2\theta} + u_{12}(u_{1\theta}\lambda_1 + u_{2\theta}\lambda_2)). \quad (1.23)$$

(1.23) provides a system of 2 quadratic equations in 2 unknowns. The solution to such a problem, if one exists, may have up to four possible roots. Solving (1.23) for λ_1 as a function of λ_2 , we obtain two functions representing the two roots from the quadratic formula: $\lambda_1^-(\lambda_2)$ and $\lambda_1^+(\lambda_2)$. We can obtain similar functions for λ_2 . The four possible roots correspond to the four possible fixed points which may exist with these functions. Two of these solutions have zero components, $\{(0, -\frac{u_{1\theta}}{u_{12}}), (-\frac{u_{2\theta}}{u_{12}}, 0)\}$, and result because we rightly assumed $\lambda_i > 0$ when we simplified (1.18). The two remaining candidates consist of the fixed points in $(\lambda_1^-, \lambda_2^-)$ and $(\lambda_1^+, \lambda_2^+)$. It is straightforward to verify that the latter pair of functions map to a set which violates (1.13)-(1.14). We must show that $(\lambda_1^-, \lambda_2^-)$ has a fixed point with the desired properties.

An examination of $(\lambda_1^-, \lambda_2^-)$ indicates that

$$(\lambda_1^-(\cdot), \lambda_2^-(\cdot)) : \quad [0, -\frac{u_{1\theta}}{u_{12}}] \times [0, -\frac{u_{2\theta}}{u_{12}}] \quad \mapsto \quad [0, -(1+\gamma)\frac{u_{1\theta}}{v_{11} + u_{11}}] \times [0, -(1+\gamma)\frac{u_{2\theta}}{v_{22} + u_{22}}],$$

and such a function is continuous. By assumption, the range is contained in the domain, and so we may apply Brouwer's theorem to establish the existence of a fixed point. Such a solution satisfies (1.13)-(1.14) and so it is incentive compatible. Moreover, it is straightforward to check that the fixed point consists of a strictly positive solution.

Next, we must check that a principal does not find it desirable to induce bunching at the corner of her rival's contract. By assumption, the conditions of A.5 are met, so bunching at a corner is not optimal.

Finally, we must show that the agent prefers the common agency environment, and the principals prefer the cooperative outcome. This follows from Corollary 6. \square

References

- ANDERSON, E. AND D. SCHMITTLEIN, 1984, "Integration of the Sales Force: An Empirical Examination," *Rand Journal of Economics*, Vol. 15, pp. 385-395.
- BARON, D., 1985, "Noncooperative Regulation of a Nonlocalized Externality," *Rand Journal of Economics*, Vol. 16, pp. 553-568.
- BERNHEIM, D. AND M. WHINSTON, 1985, "Common Marketing Agency as a Device for Facilitating Collusion," *Rand Journal of Economics*, Vol. 16, pp. 269-281.
- BERNHEIM, D. AND M. WHINSTON, 1986, "Common Agency," *Econometrica*, Vol. 54, pp. 923-943.
- BRAVERMAN, A. AND J. STIGLITZ 1982, "Sharecropping and the Interlinking of Agrarian Markets," *American Economic Review*, Vol. 72, pp. 695-715.
- CAILLAUD, B., B. JULLIEN, AND P. PICARD 1990, "Publicly Announced Contracts, Private Renegotiations, and Precommitment Effects," *mimeo*.
- DEMSKI, J. AND D. SAPPINGTON, 1984, "Optimal Incentive Contracts with Multiple Agents," *Journal of Economic Theory*, Vol. 33, pp. 152-171.
- DEMSKI, J., D. SAPPINGTON, AND P. SPILLER, 1988, "Incentive Schemes with Multiple Agents and Bankruptcy Constraints," *Journal of Economic Theory*, Vol. 44, pp. 156-167.
- FUDENBERG, D. AND J. TIROLE, 1991, *Game Theory*, forthcoming, M.I.T. Press.
- GAL-OR, E., 1989, "A Common Agency with Incomplete Information," *mimeo*, University of Pittsburgh.
- GUESNERIE, R. AND J.-J. LAFFONT, 1984, "A Complete Solution to a Class of Principal-Agent Problems with an Application to the Control of a Self-managed Firm," *Journal of Public Economics*, Vol. 25, pp. 329-369.
- HOLMSTRÖM, B. AND P. MILGROM, 1990, "Multi-Task Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design," *mimeo*, Stanford University.
- HUREWICZ, W., 1958, *Lectures on Ordinary Differential Equations*, M.I.T. Press and John Wiley & Sons, Inc.
- LAFFONT, J.-J., E. MASKIN, AND J.-C. ROCHET, 1987, "Optimal Nonlinear Pricing with Two-Dimensional Characteristics," *Information, Incentives, and Economic Mechanisms* (T. Groves, R. Radner, and S. Reiter, Eds.), pp. 256-266, University of Minnesota Press, Minneapolis.
- LAFFONT, J.-J. AND J. TIROLE, 1986, "Using Cost Observation to Regulate Firms," *Journal of Political Economy*, Vol. 94, pp. 614-641.

- LAFFONT, J.-J. AND J. TIROLE, 1989, "The Role of Government Decision-Making: A Theory of Regulatory Capture," *Quarterly Journal of Economics*, forthcoming.
- LAFFONT, J.-J. AND J. TIROLE, 1990, "Privitization and Incentives," *mimeo*, Massachusetts Institute of Technology.
- MA, C.-T., J. MOORE, AND S. TURNBULL, 1988, "Stopping Agents from Cheating," *Journal of Economic Theory*, Vol. 46, pp. 355-372.
- MARTIMORT, D., 1991, "Multi-Principaux avec Selection Adverse," *mimeo*, Institut d'Economie Industrielle, Toulouse, France, March 1991.
- MASKIN, E. AND J. TIROLE, 1990a, "The Principal-Agent Relationship with an Informed Principal: The Case of Private Values," *Econometrica*, Vol. 58, pp. 379-409.
- MASKIN, E. AND J. TIROLE, 1990b, "The Principal-Agent Relationship with an Informed Principal, II: Common Values," *mimeo*, Massachusetts Institute of Technology.
- MASTEN, S., 1988, "A Legal Basis for the Firm," *Journal of Law, Economics, and Organization*, Vol. 4, pp. 181-198.
- MCAFFEE, P. AND J. MCMILLAN, 1988, "Multidimensional Incentive Compatibility and Mechanism Design," *Journal of Economic Theory*, Vol. 46, pp. 335-354.
- MIRPLEES, J., 1971, "An Exploration in the Theory of Optimum Income Taxation," *Review of Economic Studies*, Vol. 38, pp. 175-208.
- MYERSON, R., 1981, "Optimal Auction Design," *Mathematics of Operations Research*, Vol. 6, pp. 58-73.
- MYERSON, R., 1983, "Mechanism Design by an Informed Principal," *Econometrica*, Vol. 51, pp. 1767-1797.
- OREN, S., SMITH, S., AND R. WILSON, 1983, "Competitive Nonlinear Tariffs," *Journal of Economic Theory*, Vol. 29, pp. 49-71.
- ROCHET, J.-C., 1985, "The Taxation Principle and Multi-Time Hamilton-Jacobi Equations," *Journal of Mathematical Economics*, Vol. 14, pp. 113-128.
- STIGLITZ, J. 1985, "Credit Markets and the Control of Capital," *Journal of Money, Credit, and Banking*, Vol. 17, pp. 133-152.
- TIROLE, J., 1988, *The Theory of Industrial Organization*, M.I.T. Press.
- VILLAS-BOAS, M., 1990, "Information Transfers in Vertical Relationships," *mimeo*, Massachusetts Institute of Technology.
- WILLIAMSON, O., 1985, *The Economic Institutions of Capitalism*, The Free Press.

Chapter 2

The Economics of Liquidated Damage Clauses in Contractual Environments with Private Information

Nowhere is the baneful effect of the division into specialisms more evident than in the two oldest of these disciplines, economics and law. ... [T]he rules of just conduct which the lawyer studies serve a kind of order of the character of which the lawyer is largely ignorant; and this order is studied chiefly by the economist who in turn is similarly ignorant of the character of the rules of conduct on which the order that he studies rests.

– *F. A. Hayek*

When you go to buy, don't show your silver.

– *Chinese Proverb*

2.1 Introduction

Economists have long recognized that agreements freely entered into by all effected parties with full information and cognizance of the terms of trade necessarily improve social welfare in the traditional Pareto sense. It comes as no surprise that economists look at the law with skepticism whenever courts invalidate mutually agreed upon terms

within a contract. Nonetheless, courts have routinely decided to invalidate contractually stipulated damages for breach of contract (commonly known as liquidated damages) when such damages are “unreasonably large” relative to actual or expected losses, but not those that are unreasonably small.¹ Even the courts themselves often do not know why they do what they do.

[T]he ablest of judges have declared that they felt themselves embarrassed in ascertaining the principle on which the decisions [distinguishing penalties from liquidated damages] were founded. Cotheal v. Talmadge, 9 N.Y. 551, 553 (1854).

The invalidation of excessive stipulated damage clauses is difficult to justify economically. Liquidated damage clauses promote efficiency in contractual relationships by reducing the litigation and judicial costs which accompany breach, by providing the correct incentives for a breaching party, and by optimally allocating risk.² Most importantly, stipulation of damages by the parties rather than by judicial determination allows parties to efficiently utilize their superior information which frequently courts can only imperfectly access.

The courts have had difficulty motivating the invalidation of excessive stipulated damage clauses as penalties. One theory often presented by legal scholars posits that

¹See, *Uniform Commercial Code*, §§2-302(1), 2-718(1), and the *Restatement of Contracts (Second)*, §§208, 356. The *U.C.C.*, §2-718(1) maintains:

Damages for breach by either party may be liquidated in the agreement but only at an amount which is reasonable in the light of the anticipated or actual harm caused by the breach, the difficulties of proof of loss, and the inconvenience or nonfeasibility of otherwise obtaining an adequate remedy. A term fixing unreasonably large liquidated damages is void as a penalty.

The *Restatement of Contracts (Second)*, §356(1), similarly maintains:

Damages for breach by either party may be liquidated in the agreement but only for an amount that is reasonable in the light of the anticipated or actual loss caused by the breach and the difficulties of the proof of loss. A term fixing unreasonably large liquidated damages is unenforceable on grounds of public policy as a penalty.

²Shavell [1980] analyzes the use of damage remedies to provide incentives for efficient breach. Although Shavell does not explicitly entertain the idea of stipulated damages, his analysis is closely related. In complementary work, Polinsky [1983] has shown that in some instances it is efficient from a risk-allocation viewpoint to contract for stipulated damages in excess of the actual loss from breach. Such conditions require, among other things, that the buyer should bear some of the price risk introduced from third-party, breach-inducing offers. Rea [1984, p.154], however, has argued that these conditions are rare.

legal remedies for breach of contract serve only to compensate and never to punish.³ Such a principle has economic merit. We ordinarily want parties to breach contracts when it is economically efficient that they do so. By making the promisor more than compensate the loss incurred from his nonperformance, the contract induces a suboptimal level of breach.

Unfortunately, this simple explanation falls short on two points. First, why would rational individuals agree to such a contract when there exists another contract that sets damages at the value of performance which makes both parties better off? Second, why do the courts fail to extend this operating principle to situations of under-compensatory stipulated damage agreements which produce a super-optimal level of breach?

Many courts and legal scholars answer the first question by arguing that excessive liquidated damages are presumptive evidence of a contractual failure such as fraud or mutual mistake. Arguably, courts view excessive damages as evidence that at least one party has wrongly agreed to a contract that is not Pareto improving,⁴ and respond by striking such clauses. But this does not explain why the court does not also strike extremely low liquidated damage clauses which presumably are also the product of contractual failures. We are left with an anomaly.

This paper provides an explanation for the lack of legal symmetry: While excessive damages may arguably suggest a contractual failure, undercompensatory damages are the likely result of the rational decision of two individuals bargaining in an environ-

³Farnsworth [1982, p.896] has indicated in his treatise on contract law such a principle of compensation:

If ... the stipulated sum is significantly larger than the amount required to compensate the injured party for his loss, the stipulation may have a quite different advantage to him – an *in terrorem* effect on the other party that will deter breach by compelling him to perform. Enforcement of such a provision would allow the parties to depart from the fundamental principle that the law's goal on breach of contract is not to deter breach by compelling the promisor to perform, but rather to redress breach by compensating the promisee. It is this departure that is proscribed when a court characterizes such a provision as a penalty.

⁴Aghion-Bolton [1987] provide an additional story. Two individuals may desire to sign a contract which assigns excessive liquidated damages for breach so as to foreclose entry by another supplier. Of course, these damages are socially inefficient. In a related paper, Diamond-Maskin [1979] consider the joint problems of breach and search for new trading partners. They find that because an individual who breaches can get his new partner to share the burden of the liquidated damage he pays to his old partner, a pair of partners in a contract exerts some monopoly power over potential partners, thereby making liquidated damages supercompensatory.

ment where each possesses private information about the exchange. Because the low damages do not necessarily represent a contractual failure but can realistically reflect a jointly beneficial contract arrived at under the constraints of asymmetric information, legal institutions are arguably correct in enforcing such terms unless there exists other hard evidence of contractual failures such as fraud or mutual mistake. This essay's principle thesis maintains that when each party to a contract possesses private information whose disclosure would adversely affect its position in the contractual bargaining, rationally calculated liquidated damages will be set at *under-compensatory* levels.

Current economic analysis of liquidated damage clauses has been limited to symmetrically informed parties. In many contractual situations, however, the assumption that parties entered into the contract without private information is not palatable. When such asymmetries in information are present, the liquidated damage clause takes on dual roles: (i) providing incentives for efficient breach, and (ii) efficiently screening among different types of buyers and sellers. Specifically, this paper demonstrates that when parties have asymmetric information, stipulated damages may be used to communicate valuable information at the pre-contractual stage. As such, the loss from insufficient or excessive breach may be offset by informational gains. In fact, in the typical buyer-seller contract where each party has private information, stipulated damages will almost always fall short of actual losses from the breach.

This paper examines the buyer-seller relationship, although its results appear much more general. We assume that the buyer has private information regarding the value of the product to herself, and that the seller has private information regarding alternative markets where the product may be sold absent a sale to the present buyer. The contractual framework is modeled in Section 2. In Section 3 we examine various bargaining situations. In section 3.1, we analyze the consequences of placing all of the bargaining power in the hands of the buyer; in Section 3.2, we assume all of the bargaining power resides with the seller.⁵ Later, in Section 3.3, we examine what an efficient arbitrator would assign as stipulated damages. In both one-sided bargaining and the arbitration sce-

⁵Placing all of the bargaining power in the hands of one party manifests itself as the opportunity of the party to write a contract and make a take-it-or-leave-it offer to the other.

narios, we find that there is no role for excessive stipulations, but there is a positive role for under-compensatory terms. Indeed, under-compensatory terms occur with probability one. These terms provide a valuable method for both parties to signal to the other their private information, increasing the gains from trade.

Section 4 examines the policy question of whether a perfectly informed court would generally improve matters by requiring that all stipulated damages be exactly compensating. We find that under plausible conditions, even a perfectly informed court can be a menace to the parties' contract and to social welfare if it naively imposes a requirement that liquidated damages equal actual value, *ex post*. There is a direct benefit and a direct cost from judicial intervention. Eliminating the agents' abilities to set liquidated damages below valuation reduces inefficient breach of contract. If $\ell = v$, the seller will breach only if it is efficient to do so. Unfortunately, such a restriction on liquidated damages also restricts the offerer to contracts which set a single price. This restriction may lead to buyer-designed contracts which only induce trade with low-opportunity sellers, and seller-designed contracts which only induce trade with high-value buyers. Consequently, some individuals may be foreclosed from trade, leaving unrealized gains from the exchange. These results are analogous to the social planner's decision of whether to allow second-degree price discrimination in the context of monopoly pricing. Section 5 summarizes and concludes.

2.2 The Contractual Framework

We examine the contractual relationship between a buyer and a seller, where third-party offers for the seller's services may induce breach after an agreement has been reached. The buyer and seller recognize this possibility and bargain both over price and a damage stipulation which the seller agrees to pay the buyer in event of non-performance.⁶

The buyer (she) and a seller (he) contract to trade a single good at date 1. After contracting, the buyer cannot find other sellers (e.g., the buyer makes relation-specific

⁶The character of the results remain unchanged if one considers instead that the buyer may breach after finding an alternative product. In this alternative, both parties negotiate damages which the buyer will compensate the seller for the lost transaction. The issue of robustness is briefly discussed in Section 7.

investments or her outside opportunities disappear) but the seller's opportunism is constrained by the non-performance damage terms of the contract. At date 2 a third-party offer is made to the seller for his wares. The seller can either accept the third-party's offer and pay the buyer the stipulated damages, or deliver the product to the buyer.

The buyer has valuation, v , distributed according to the continuous, positive density function, $f(\cdot)$, on $[\underline{v}, \bar{v}]$, with cumulative distribution, $F(\cdot)$. Only the buyer knows v , although its distribution is common knowledge. The third-party's offer for the seller's product is equal to $\theta + \epsilon$, where θ is known by the seller at date 1, and ϵ is an unknown outside valuation shock at date 2. θ is distributed according to the continuous, positive density function, $g(\theta)$ on $[\underline{\theta}, \bar{\theta}]$, with cumulative distribution, $G(\theta)$, such that $\bar{v} > \underline{\theta}$ (i.e., there are gains from trade with some probability). Only the seller knows θ , although its distribution is common knowledge. ϵ is distributed according to the continuous, positive density function, $h(\epsilon)$, on $[\underline{\epsilon}, \bar{\epsilon}]$, with cumulative distribution, $H(\epsilon)$.⁷ Neither party observes ϵ at date 1, and only the seller observes ϵ at date 2. The expected value of ϵ is zero, thereby making θ an unbiased estimate at date 1 of the alternative market at date 2. Additionally, it is common knowledge that the seller's costs are to be zero, although zero costs is without loss of generality. Absent any contract offer, the seller expects to make θ . The buyer's outside opportunities have been normalized to zero. Finally, I assume that $\frac{1-F(v)}{f(v)}$ is nonincreasing in v , $\frac{G(\theta)}{g(\theta)}$ is nondecreasing in θ , and payoffs are not discounted.

⁷To ensure the global concavity of each of the several maximization programs in this paper, we assume that for all θ and for all $\ell \in [\theta + \underline{\epsilon}, \theta + \bar{\epsilon}]$,

$$\begin{aligned}
 h(\ell - \theta) &> h'(\ell - \theta)[v - \ell], \\
 h(\ell - \theta) &> h'(\ell - \theta)[v - \ell - \frac{G(\theta)}{g(\theta)}], \\
 h(\ell - \theta) &> h'(\ell - \theta)[v - \ell - \frac{1 - F(v)}{f(v)}], \\
 h(\ell - \theta) &> h'(\ell - \theta)[v - \ell - \frac{G(\theta)}{g(\theta)} - \frac{1 - F(v)}{f(v)}].
 \end{aligned}$$

These conditions are made for tractability; a weaker (but more complicated) set of conditions would also suffice. In any case, if h is a uniform distribution, these conditions are trivially satisfied. Additionally, we further assume that the support of ϵ is sufficiently large so as to eliminate corner solution problems. The latter condition is also for simplicity and would be satisfied, for example, if $F(v)$, $G(\theta)$, and $H(\epsilon)$ are uniform distributions and $\underline{\epsilon} \leq \min\{\underline{\theta}, \underline{v} + \underline{\theta} - 2\bar{\theta}, 2\underline{v} - \bar{\theta} - \bar{v}\}$ and $\bar{\epsilon} \geq \bar{v} - \underline{\theta}$.

The contract consists of a binary decision to agree to trade, δ ($\delta = 1$ if there is an agreement to trade, $\delta = 0$, otherwise), a price, p , paid at the time of signing, and a stipulated damage payment of ℓ to be paid at date 2 in the event of the supplier's breach after a decision to trade has been made. Thus, a contract outcome is given by $\{\delta, p, \ell\}$. As is standard, we restrict our attention to deterministic, piecewise C^1 contracts. For now, we assume that only the contract and the existence of breach is observable by the court. Later in Section 4 we relax this assumption to determine if a perfectly informed, but myopic, court could always improve contracting among the parties.

Given ℓ , the supplier will breach whenever $\theta + \epsilon > \ell$, and perform otherwise. Thus, the probability of performance is $H(\ell - \theta)$ and the probability of breach is $1 - H(\ell - \theta)$.

The *net* profit of the supplier from a contract, $\{\delta, p, \ell\}$, is

$$\pi^s(\theta) = \delta \left(p + \int_{\ell - \theta}^{\bar{\epsilon}} [\theta + \epsilon - \ell] dH(\epsilon) - \theta \right),$$

where the second term in the parentheses represents the expected gain from breach when the outside opportunity is lucrative, and the third term in parentheses is the seller's opportunity cost in agreeing to a contract (i.e., the lost expected profit, θ). The profit of the buyer is

$$\pi^b(v, \theta) = \delta [vH(\ell - \theta) + \ell[1 - H(\ell - \theta)] - p].$$

I consider three different contracting scenarios to provide a range of environments for analysis. First, the buyer may propose the contract to the seller, and the seller may accept or reject it. Second, the seller may propose the contract, and the buyer may accept or reject it. Finally, an uninformed third party may design a contract which maximizes the joint surplus from trade between the parties.

Before considering each case in the following sections, we consider the full-information benchmark solutions for comparisons: In all three cases, the optimal full-information contract involves trade (i.e., $\delta = 1$) if and only if

$$(v - \theta)H(v - \theta) + \int_{v - \theta}^{\bar{\epsilon}} (\theta + \epsilon - v) dH(\epsilon) \geq 0,$$

and in such case, $\ell = v$. The above expression represents the expected gains from trade given that $\ell = v$. The first term is the expected gain under performance of the exchange; the second term is the option value of the outside opportunity that is available to the seller whenever $\theta + \epsilon \geq v$.

If the buyer has all of the bargaining power, the buyer's optimal strategy is to maximize her profits subject to the seller's acceptance of the conditions (i.e., $\pi^* \geq 0$). Substituting for p and simplifying yields the following program for the solution of δ and ℓ :

$$\max_{\delta, \ell} \delta \left(vH(\ell - \theta) + \ell[1 - H(\ell - \theta)] - \theta + \int_{\ell - \theta}^{\bar{\epsilon}} (\theta + \epsilon - \ell) dH(\epsilon) \right). \quad (2.1)$$

The necessary first-order condition is $\ell = v$. Given our assumptions on $h(\epsilon)$, this is also sufficient. $\delta = 1$ whenever (2.1) is positive at $\ell = v$. The contract price offered by the buyer is

$$p = \theta - \int_{v - \theta}^{\bar{\epsilon}} (\theta + \epsilon - v) dH(\epsilon). \quad (2.2)$$

Similarly, if the seller has all of the bargaining power, the seller's optimal strategy is to maximize his profits, subject to the buyer's acceptance of terms (i.e., $\pi \geq 0$). Substituting for p and simplifying yields

$$\max_{\delta, \ell} \delta \left(vH(\ell - \theta) + \ell[1 - H(\ell - \theta)] - \theta + \int_{\ell - \theta}^{\bar{\epsilon}} (\theta + \epsilon - \ell) dH(\epsilon) \right).$$

which is identical to the buyer's program above for the choice of δ and ℓ . Thus, we again find $\ell = v$. Note, however, the price paid by the buyer to the seller under this scheme is $p = v$, which extracts all of the buyer's rent. Finally, if an uninformed arbiter proposes a contract for the parties, the arbiter will maximize the expected gains from trade by choosing ℓ to maximize the collective surplus

$$\int_{\underline{\theta}}^{\bar{\theta}} \int_{\underline{v}}^{\bar{v}} \left\{ (v - \theta)H(\ell - \theta) + \int_{\ell - \theta}^{\bar{\epsilon}} \epsilon dH(\epsilon) \right\} dF(v) dG(\theta). \quad (2.3)$$

Again, the solution is to set $\ell = v$. The third party then chooses a price to split the gains from trade with $p \in \left[\theta - \int_{v - \theta}^{\bar{\epsilon}} (\theta + \epsilon - v) dH(\epsilon), v \right]$. It is not surprising that the optimal full-information contract under trade specifies $\ell = v$ for each contracting environment,

since this condition guarantees that breach occurs if and only if it is efficient.

When information is not public, the resulting contract typically has $\ell \neq v$. Instead, ℓ will depend upon v and θ in a manner which will elicit a party's private information by creating distortions from efficient breach. The precise relationship between ℓ , v , and θ will depend fundamentally on the contractual context: Buyer power, Seller power, or Third-party Arbitration.

2.3 Contracting Environments

2.3.1 The Buyer's Optimal Contract

Because the buyer does not know the seller's expected outside opportunity, θ , she must take into account the effect of the liquidated damage clause on the seller's gains from trade. If ℓ is set arbitrarily high, it will effectively lock the seller out of the alternative market; a low ℓ preserves the option value of breach, which in turn is an increasing function of θ . For seller's with high θ 's, this will require a higher price to offset the loss in opportunity. Recognizing this relationship, the buyer can effectively use the damage clause to screen among different types of sellers much in the same way that a price-discriminating monopolist screens among different consumers by offering multiple quantity-price packages. The buyer will offer a menu of contracts, from which the seller chooses the one most profitable given his θ .⁸ That is, the buyer may offer a continuum of contracts to the seller represented by a function $p(\ell)$. Following the Revelation Principle (see e.g., Myerson [1985] or Fudenberg-Tirole [1991]), we reparameterize according to the seller's outside opportunity, θ , as $x(\theta) \equiv \{\delta(\theta), p(\theta), \ell(\theta)\}$. Accordingly, we may solve for the buyer's optimal choice of x for every θ , subject to each seller type finding it optimal to choose the contract designed for his type.

⁸The choice of contract by the buyer may possibly reveal information about the buyer's type, v , to the seller. There is no problem with mechanism design by informed principal in this case, however, as the seller's utility is independent of v , and the buyer has no action to take which could indirectly affect the seller's utility.

The buyer's expected profit from any mechanism $x(\cdot)$ is

$$\pi^b(v) = \int_{\underline{\theta}}^{\bar{\theta}} \delta(\theta) \{vH(\ell(\theta) - \theta) + \ell(\theta)[1 - H(\ell(\theta) - \theta)] - p(\theta)\} dG(\theta). \quad (2.4)$$

She maximizes this subject to two sets of constraints: the seller must be willing to sign the contract (i.e, not make a loss from trade) and the seller must select the contract designed for his type.

Define $\pi^s(\hat{\theta}|\theta)$ as the profit to a seller with outside opportunity θ , who selects the contract designed for a seller of type $\hat{\theta}$. That is,

$$\pi^s(\hat{\theta}|\theta) \equiv \delta(\hat{\theta}) \left(p(\hat{\theta}) + \int_{\ell(\hat{\theta})-\theta}^{\bar{\epsilon}} (\theta + \epsilon - \ell(\hat{\theta})) dH(\epsilon) - \theta \right). \quad (2.5)$$

The buyer's first constraint requires that every seller's truthful selection must yield non-negative profits. Thus,

$$\pi^s(\theta|\theta) \geq 0 \quad (2.6)$$

for all θ , which represents the *individual rationality* (IR) or participation constraint. Second, every seller must select the correct contract from the menu. These *incentive compatibility* (IC) constraints require

$$\pi^s(\theta|\theta) \geq \pi^s(\hat{\theta}|\theta), \forall \theta, \hat{\theta}. \quad (2.7)$$

The IR and IC constraints in (2.6) and (2.7) are intractable in their present form so we follow the standard procedure of replacing them with the significantly simpler representation in Lemma 1.

Lemma 1 *The mechanism $\{p(\theta), \ell(\theta)\}$ satisfies the IR and IC constraints if*

$$\pi^s(\theta) = \pi^s(\bar{\theta}) + \int_{\theta}^{\bar{\theta}} \delta(t)H(\ell(t) - t)dt, \quad (2.8)$$

$$\pi^s(\bar{\theta}) \geq 0, \quad (2.9)$$

$$\delta(\theta)H(\ell(\theta) - \theta) \geq \delta(\theta')H(\ell(\theta') - \theta), \forall \theta' > \theta. \quad (2.10)$$

In addition, (2.8) and (2.9) are necessary conditions for IR and IC.

The proof is standard and provided in the Appendix. Intuitively, the lemma follows from the envelope theorem: Assuming truthful selection is optimal for the seller, totally differentiating (2.5) results in $\frac{d\pi^s(\theta|\theta)}{d\theta} = -H(\ell(\theta) - \theta)$. Integrating this derivative produces (2.8) and (2.9). The condition in (2.10) is a second-order condition for truthful selection.

With this simplification, we proceed by substituting (2.8)-(2.9) into the buyer's objective, (2.4). Equating (2.8) with $\pi^s(\theta|\theta)$ and solving for $\delta(\theta)p(\theta)$ yields

$$\delta(\theta)p(\theta) = \theta - \int_{\ell(\theta)-\theta}^{\bar{\epsilon}} \delta(\theta)(\theta + \epsilon - \ell(\theta))dH(\epsilon) + \pi^s(\bar{\theta}) + \int_{\theta}^{\bar{\theta}} \delta(t)H(\ell(t) - t)dt. \quad (2.11)$$

Recognizing that the buyer will optimally set $\pi^s(\bar{\theta}) = 0$, taking the expectation of $\delta(\theta)p(\theta)$ over θ , and integrating by parts yields

$$\int_{\underline{\theta}}^{\bar{\theta}} \delta(\theta)p(\theta)dG(\theta) = \int_{\underline{\theta}}^{\bar{\theta}} \delta(\theta) \left\{ \theta - \int_{\ell(\theta)-\theta}^{\bar{\epsilon}} (\theta + \epsilon - \ell(\theta))dH(\epsilon) + H(\ell(\theta) - \theta) \frac{G(\theta)}{g(\theta)} \right\} dG(\theta). \quad (2.12)$$

Substituting this expression into the buyer's objective function yields the unconstrained problem

$$\max_{\delta, \ell} \int_{\underline{\theta}}^{\bar{\theta}} \delta(\theta) \left\{ \left(v - \theta - \frac{G(\theta)}{g(\theta)} \right) H(\ell(\theta) - \theta) + \int_{\ell(\theta)-\theta}^{\bar{\epsilon}} \epsilon dH(\epsilon) \right\} dG(\theta), \quad (2.13)$$

which may be solved by maximizing ℓ pointwise over θ and checking that the solution satisfies (2.10). This program yields the following Proposition.

Proposition 1 *The optimal menu of contracts, $\{\delta(\theta), p(\theta), \ell(\theta)\}$, for the buyer consists of a contract with $\ell(\theta) = v - \frac{G(\theta)}{g(\theta)}$, $p(\theta)$ such that (2.11) is satisfied, and $\delta(\theta)$ such that*

$$\delta(\theta) = \begin{cases} 1 & \forall \theta \in [\underline{\theta}, \theta^*) \\ 0 & \forall \theta \in [\theta^*, \bar{\theta}] \end{cases}$$

where θ^* is either the unique value of $\theta \in [\underline{\theta}, \bar{\theta}]$ such that the integrand in (2.13) is zero, if such a value exists, or $\theta^* = \bar{\theta}$ otherwise.

Proof: Ignoring the decision to trade, pointwise maximization of (2.13) yields the expression for $\ell(\theta)$, which is monotonically decreasing. (Given our assumptions on H , F , and G , this pointwise optimization program is concave in ℓ . (2.11) provides us with $p(\theta)$ such that IC and IR are satisfied if the monotonicity condition in (2.10) is satisfied. Pointwise maximization of $\delta(\theta)$ yields $\delta = 1$ whenever the integrand in (2.13) is non-negative. Given our assumptions on the inverse hazard rate, the integrand in (2.13) is strictly decreasing in θ , which implies that $\delta(\theta)$ is decreasing in θ and that there is at most one value of θ such that the integrand in (2.13) is exactly zero; thus θ^* is unique. Consequently, the monotonicity condition in (2.9) are satisfied and so the mechanism is IC and IR. \square

As the Proposition indicates, the actual buyer loss from breach, v , almost always exceeds the amount of stipulated damages in the optimal contract when the buyer has all of the bargaining power.

2.3.2 The Seller's Optimal Contract

Because the seller does not know the buyer's valuation of the good, he must take into account the effect of the liquidated damage clause on protecting the buyer's value. If ℓ is set arbitrarily low, it will allow the seller to breach and use the alternative market whenever ϵ is favorable, thereby imposing a loss of v on the buyer. For buyers with high v 's, this will produce a lower reserve price due to the lower likelihood that the value of the bargain will accrue. Recognizing this, the seller can effectively use the stipulated damage clause to select among the different types of buyers just as the buyer was previously shown to select among sellers.

The seller may offer a menu of contracts like that in the previous section and allow the buyer to choose the one most profitable given her v . In this case, the menu can be represented by either the function $p(\ell)$ or the parametric triplet $y(v) \equiv \{\delta(v), p(v), \ell(v)\}$.

The seller's expected profit from any mechanism $y(\cdot)$ is

$$\pi^s(\theta) = \int_{\underline{v}}^{\bar{v}} \delta(v) \left\{ p(v) + \int_{\ell(v)-\theta}^{\bar{c}} (\theta + \epsilon - \ell(v)) dH(\epsilon) - \theta \right\} dF(v). \quad (2.14)$$

He maximizes this subject to the buyer's IR and IC constraints, analogous to the Buyer's problem above. In this section for simplicity we also assume that θ is observed by the buyer.⁹

Define $\pi^b(\hat{v}|v)$ by

$$\pi^b(\hat{v}|v) \equiv vH(\ell(\hat{v}) - \theta) + \ell(\hat{v})[1 - H(\ell(\hat{v}) - \theta)] - p(\hat{v}),$$

which represents the profit to a seller with outside opportunity v who selects the contract designed for a seller of type \hat{v} . Analogously to the buyer-contract case, the IR and IC constraints are, respectively,

$$\pi^b(v|v) \geq 0,$$

$$\pi^b(v|v) \geq \pi^b(\hat{v}|v),$$

for all v and \hat{v} . Again, as in the buyer-designed contract, the above two sets of constraints are difficult to work with but can be greatly simplified, as in Lemma 2.

Lemma 2 *The mechanism $\{\delta(v), p(v), \ell(v)\}$ satisfies the buyer's IR and IC constraints if*

$$\pi^b(v) = \pi^b(\underline{v}) + \int_{\underline{v}}^v \delta(t)H(\ell(t) - \theta)dt, \quad (2.15)$$

$$\pi^b(\underline{v}) \geq 0, \quad (2.16)$$

$$\delta(v)H(\ell(v) - \theta) \geq \delta(v')H(\ell(v') - \theta), \forall v, v'. \quad (2.17)$$

⁹Because the buyer's expected returns from trade depend negatively on θ , the seller might otherwise attempt to signal to the buyer that θ is in fact low by making a contractual offer which only low- θ sellers would find profitable to make. The general problem of mechanism design by an informed principal has been studied by Maskin-Tirole [1990]. Rather than assuming θ is known by the buyer, an alternative way to avoid the informed-principal problem is by assuming that all sellers offer the same contract (i.e., they collectively pool), $\underline{\theta}$ is sufficiently large that all types of sellers find it optimal to contract with the buyer (i.e., in terms of Proposition 2, $v^* = \underline{v}$), and any seller who offers a different contract than the expected pooled contract is assumed to be a high-type θ by the buyer. Because the equilibrium contracts derived in Proposition 2 are independent of θ , this forms a Bayesian-Nash equilibrium.

In addition, (2.15)-(2.16) are necessary conditions for IR and IC.

The proof is provided in the Appendix and again is an application of the envelope theorem. With this simplification, we proceed by substituting (2.15)-(2.16) into the seller's objective, (2.14). Equating (2.15) with $\pi^b(v|v)$ and solving for $\delta(v)p(v)$ yields

$$\delta(v)p(v) = vH(\ell(v) - \theta) + \ell(v)[1 - H(\ell(v) - \theta)] - \pi^b(\underline{v}) - \int_{\underline{v}}^v \delta(t)H(\ell(t) - \theta)dt.$$

Recognizing that the seller will optimally set $\pi^b(\underline{v}) = 0$, taking the expectation of $\delta(v)p(v)$ over v , and integrating by parts, produces

$$\begin{aligned} \int_{\underline{v}}^{\bar{v}} \delta(v)p(v)dF(v) = \\ \int_{\underline{v}}^{\bar{v}} \delta(v) \left\{ vH(\ell(v) - \theta) + \ell(v)[1 - H(\ell(v) - \theta)] - H(\ell(v) - \theta) \frac{1 - F(v)}{f(v)} \right\} dF(v). \end{aligned} \quad (2.18)$$

Substituting this expression into the buyer's objective function yields the unconstrained problem

$$\max_{\delta, \ell} \int_{\underline{v}}^{\bar{v}} \delta(v) \left\{ \left(v - \theta - \frac{1 - F(v)}{f(v)} \right) H(\ell(v) - \theta) + \int_{\ell(v) - \theta}^{\bar{\varepsilon}} s dH(\varepsilon) \right\} dF(v). \quad (2.19)$$

This problem may be solved by maximizing ℓ pointwise over v , and checking that the resulting solution satisfies (2.17). The resulting expressions provide Proposition 2. As the Proposition indicates, ℓ is again almost always below actual loss.

Proposition 2 *The optimal menu of contracts, $\{\delta(v), p(v), \ell(v)\}$, for the seller sets $\ell(v) = v - \frac{1 - F(v)}{f(v)}$ and $p(v)$ such that (2.18) holds above. Additionally, the seller chooses,*

$$\delta(v) = \begin{cases} 1 & \forall v \in (v^*, \bar{v}] \\ 0 & \forall v \in [\underline{v}, v^*] \end{cases}$$

where v^* is either the unique value of $v \in [\underline{v}, \bar{v}]$ such that the integrand in (2.19) is zero, if such a value exists, or $v^* = \underline{v}$ otherwise.

Proof: Ignoring the decision to trade, pointwise maximization of (2.19) yields the expression for $\ell(v)$, which is monotonically increasing. (Given our assumptions on F , G , and H , the pointwise optimization program is concave in ℓ .) (2.18) provides us with $p(v)$ such that IC and IR are satisfied if the monotonicity condition in (2.17) is satisfied. Pointwise maximization of $\delta(v)$ yields $\delta = 1$ whenever the integrand in (2.19) is nonnegative. Given our assumptions on the inverse hazard rate, the integrand in (2.19) is strictly increasing in v , which implies that $\delta(v)$ is decreasing in v and that there is at most one value of v such that the integrand in (2.19) is exactly zero; thus v is unique. Consequently, the monotonicity conditions in (2.17) are satisfied and so the mechanism is IC and IR. □

2.3.3 Brokered Contracts

Rather than place all of the bargaining power in the hands of one agent, we now consider the resulting contract where both agents delegate the contractual terms to a third-party (e.g., a broker or arbiter) who knows neither v nor θ .¹⁰ This broker is concerned only with maximizing the total gains from trade when each party knows only its own private information.¹¹ Before, when one party had full contractual power, that party traded off breach inefficiencies against increased rent extraction. Under such a skewed bargaining environment, ℓ never exceeds v in the optimal contract. We now find that even when a broker is employed, the optimal contract never involves excessive stipulated damages.

The problem facing the broker is to maximize the joint gains from trade by designing a menu of contracts. The contracts may depend upon both θ and v . We can think of the menu as an offer of a menu of menus to the buyer, one of which the buyer selects. From

¹⁰In this section, we return to our assumption that the buyer does not observe θ .

¹¹As a motivation, one might suppose that certain institutions evolve which maximize the joint gains from trade between agents from an ex ante point of view, and agents use these institutions in order to avoiding signaling adverse information to one another, although this motivation is admittedly very loose. An alternative motivation has the buyer and seller contracting ex ante, before they learn their private information, but subject to a limited liability constraint where either party can legally walk away from the contract once private information is learned if losses are sufficiently great. This latter explanation may be realistic in the requirements contracting context.

the selected menu, the seller is allowed to choose the final contract. Alternatively, we may parameterize this family of contracts by (θ, v) , and envision the contract as a direct revelation mechanism where each party announces his or her private information and the broker selects the appropriate contract according to $z(\theta, v) \equiv \{\delta(\theta, v), p(\theta, v), \ell(\theta, v)\}$.

Because there is two-sided asymmetric information, the traditional techniques need to be augmented slightly; we follow Myerson-Satterthwaite [1983] in this regard.¹² Lemma 3 is a direct extension of Lemmas 1 and 2 and characterizes the set of all contracts which are incentive compatible and individually rational for both parties. Proposition 3 provides the solution of the broker's problem.

Lemma 3 *The mechanism $\{\delta(\theta, v), p(\theta, v), \ell(\theta, v)\}$ satisfies IR and IC constraints if and only if*

$$\pi^s(\theta) = \int_{\underline{v}}^{\bar{v}} \left\{ \pi^s(\bar{\theta}, v) + \int_{\underline{\theta}}^{\bar{\theta}} \delta(\theta, v) H(\ell(s, v) - s) ds \right\} dF(v), \quad (2.20)$$

$$\pi^b(v) = \int_{\underline{\theta}}^{\bar{\theta}} \left\{ \pi^b(\theta, \underline{v}) + \int_{\underline{v}}^v \delta(\theta, v) H(\ell(\theta, t) - \theta) dt \right\} dG(\theta), \quad (2.21)$$

$$\pi^s(\bar{\theta}) \geq 0, \quad (2.22)$$

$$\pi^b(\underline{v}) \geq 0, \quad (2.23)$$

$$\int_{\underline{v}}^{\bar{v}} \delta(\theta, v) H(\ell(\theta, v) - \theta) dF(v) \leq \int_{\underline{v}}^{\bar{v}} \delta(\theta', v) H(\ell(\theta', v) - \theta') dF(v), \forall \theta > \theta', \quad (2.24)$$

$$\int_{\underline{\theta}}^{\bar{\theta}} \delta(\theta, v) H(\ell(\theta, v) - \theta) dG(\theta) \geq \int_{\underline{\theta}}^{\bar{\theta}} \delta(\theta, v') H(\ell(\theta, v') - \theta) dG(\theta), \forall v > v'. \quad (2.25)$$

Alternatively, there exists a function $p(\theta, v)$ such that $\ell(\theta, v)$ is IC and IR if and only if (2.24) and (2.25) hold, and

$$\int_{\underline{\theta}}^{\bar{\theta}} \int_{\underline{v}}^{\bar{v}} \left\{ \left[\left(v - \frac{1 - F(v)}{f(v)} \right) - \left(\theta - \frac{G(\theta)}{g(\theta)} \right) \right] H(\ell(\theta, v) - \theta) + \int_{\ell(\theta, v) - \theta}^{\bar{\epsilon}} \epsilon dH(\epsilon) \right\} dF(v) dG(\theta) \geq 0. \quad (2.26)$$

¹²Also see Williams [1987] for a fuller treatment and extension of Myerson-Satterthwaite's model. Williams characterizes the efficient locus of contracts, depending upon the weights attached to the buyer and seller's utility.

The proof of Lemma 3 follows from Lemmas 1 and 2 and is contained in the Appendix.

With this Lemma, we may write the arbitrator's problem as maximizing the expected profit of each party subject to (2.26) above. That is

$$\max_{\delta, \ell} \int_{\underline{\theta}}^{\bar{\theta}} \int_{\underline{v}}^{\bar{v}} \delta(\theta, v) \left\{ (v - \theta)H(\ell(\theta, v) - \theta) + \int_{\ell(\theta, v) - \theta}^{\bar{v}} \epsilon dH(\epsilon) \right\} dF(v)dG(\theta), \quad (2.27)$$

subject to (2.26). Let $\mu \ell(\theta, v)$ be the Lagrange multiplier for (2.26). We multiply μ by δ without loss of generality as the IC and IR constraints do not bind when $\delta = 0$. Bringing the constraint into the integral of (2.27) and simplifying yields

$$\int_{\underline{\theta}}^{\bar{\theta}} \int_{\underline{v}}^{\bar{v}} (1 + \mu)\delta(\theta, v) \left\{ \left[(v - \theta) - \frac{\mu}{1 + \mu} \left(\frac{1 - F(v)}{f(v)} + \frac{G(\theta)}{g(\theta)} \right) \right] H(\ell(\theta, v) - \theta) + \int_{\ell(\theta, v) - \theta}^{\bar{v}} \epsilon dH(\epsilon) \right\} dF(v)dG(\theta). \quad (2.28)$$

The functions δ and ℓ which maximize this integral may be found by maximizing the expression for δ and ℓ pointwise in θ and v , and checking that the solution satisfies (2.24) and (2.25). The solution results in the following Proposition.

Proposition 3 *The optimal contracts for the arbitrated buyer-seller relationship consists of*

$$\ell(\theta, v) = v - \frac{\mu}{1 + \mu} \left(\frac{1 - F(v)}{f(v)} + \frac{G(\theta)}{g(\theta)} \right),$$

where $\mu \geq 0$, and with

$$\delta(v) = \begin{cases} 1 & \forall \theta, v, \text{ s.t. (2.28) is nonnegative} \\ 0 & \text{otherwise.} \end{cases}$$

Additionally, $\mu > 0$ if

$$\int_{\underline{\theta}}^{\bar{\theta}} \int_{\underline{v}}^{\bar{v}} \left\{ \left[\left(v - \frac{1 - F(v)}{f(v)} \right) - \left(\theta + \frac{G(\theta)}{g(\theta)} \right) \right] H(v - \theta) + \int_{v - \theta}^{\bar{v}} \epsilon dH(\epsilon) \right\} dF(v)dG(\theta) > 0.$$

Proof: Maximizing the above expression over δ and ℓ pointwise in θ and v yields $\delta(\theta, v)$

and $\ell(\theta, v)$. These in turn provide for the construction of $p(\theta, v)$ that satisfies (2.20)-(2.21). ℓ is nonnegative, and therefore $\ell(\theta, v)$ is nondecreasing in v and nonincreasing in θ . Moreover, at the optimum, (2.28) is increasing in v and decreasing in θ , implying that $\delta(\theta, v)$ is nondecreasing in v and nonincreasing in θ . Hence, (2.24)-(2.25) are satisfied and the mechanism is IC and IR. To prove that $\mu > 0$ under the integral condition above, suppose to the contrary that $\mu = 0$. Then $\ell(\theta, v) = v$, and by our hypothesis, (2.26) must fail, indicating that $\mu > 0$. □

Proposition 3 demonstrates that stipulated damages do not exceed the actual loss from breach of contract and are strictly less than actual loss whenever the expected gains from trade are less than the expected information rents for almost all θ . Because the arbitration contract yields greater combined gains from trade than either the buyer-contract or the seller-contract, our earlier results are robust to the distribution of bargaining power.

2.4 Welfare Implications and Policy Conclusions

We have seen that the existence of private information by contracting parties in a wide range of bargaining environments introduces the likelihood that liquidated damages will be below the actual losses caused by breach. Using liquidated damages to select among different types of economic agents is, in a sense, second-degree price discrimination where the monopolist offers price-damage, rather than price-quantity, bundles. The question then arises as to whether public policy should require that all damages for breach of contract equal the true losses incurred, providing such information about losses is available to the court after the breach. It is arguable that any intervention would be precarious at best, especially given our limitations of knowledge about the actual contracting conditions between parties.

To consider the issue of judicial intervention, we posit the strongest possible assumption in favor of activism to determine the most optimistic assessment: assume that courts can perfectly determine actual losses from breach *ex post*. That is, assume v becomes known to the court and to both parties at date 2 in the event of breach. With this as-

sumption, I seek to answer the question of whether the court should require $\ell = v$ in all breached contracts. For realism, further assume that parties cannot base their contract price on the judicial determination of v ; otherwise, the court would become nothing more than an auditing agency for private contracts. That is, observed v can only be used to determine ℓ .

There is a benefit and a cost from judicial intervention: Eliminating the agents' abilities to set liquidated damages below valuation reduces inefficient breach of contract, but may foreclose some buyers and sellers from efficient trade.

If there were only two possible types of buyers and sellers, and assuming the offerer would chose to serve only part of the market if it were permissible to choose $\ell \neq v$, then judicial intervention always produces inefficiencies. To see this, note that when the offeree is of a good type (i.e., high-value buyer or low-opportunity seller), liquidated damages are set at actual value. Consequently, for good types there is no inefficiency with or without judicial intervention. When the offerees are of bad type (i.e., low-value buyers or high-opportunity sellers), offerers will set inefficient damage levels when given the option. But while the terms are inefficient, individual rationality implies both parties are better off trading than not trading. Judicial intervention that prevents the use of under-compensatory damages must therefore decrease social welfare.

When a continuum of types of buyers and sellers exist, the analysis is more difficult. Consider the problem facing the buyer with all of the bargaining power who is constrained to set $\ell = v$, ex post, in all offered contracts. If she sets the contract price low, only very low θ -type sellers will accept the terms, but she will make a larger profit on those contracts where such a sale is made. If she sets the price high, her terms will be accepted by most sellers, but her gains from actual trade will be lower. The problem facing her is much the same as that facing a monopolist setting one price: higher prices result in fewer sales but greater profits per sale. Her maximization problem is simply,

$$\max_p (v - p)G(\theta(p)),$$

where $\theta(p)$ is defined as the highest θ -type seller who would be willing to buy at price

p , and is given implicitly by (2.2). Each buyer will set a different price depending on her type v , just as a monopolist's prices vary with marginal cost. Because (2.2) implies that $1/\theta'(p) = H(v - \theta(p))$, maximization reveals that the optimal constrained-contract price is implicitly given by $p = v - H(v - \theta(p)) \frac{G(\theta(p))}{g(\theta(p))}$, which defines $p^*(v)$. Using (2.2) again allows us to define $\theta^*(v)$ as the threshold type of seller who chooses not to sell at the buyer's offered price of $p^*(v)$. All sellers with types lower than $\theta^*(v)$ sell to the buyer at buyer's asking price of $p^*(v)$. That is, $\theta^*(v)$ is defined by

$$\left(v - \theta^* - \frac{G(\theta^*)}{g(\theta^*)} \right) H(v - \theta^*) + \int_{v-\theta^*}^{\bar{\epsilon}} \epsilon dH(\epsilon) = 0.$$

Consider now the seller's problem when the seller has all of the bargaining power. Because of the constraint that $\ell = v$, the seller's maximization problem is to choose p to solve

$$\max_p \int_p^v \left(p + \int_{v-\theta}^{\bar{\epsilon}} (\theta + \epsilon - v) dH(\epsilon) \right) dF(v) + \int_{\underline{v}}^p \theta dF(v).$$

Maximization reveals that the seller's optimal constrained-contract price, satisfies

$$p = \theta + \frac{1 - F(p)}{f(p)} - \int_{p-\theta}^{\bar{\epsilon}} (\theta + \epsilon - p) dH(\epsilon).$$

We can analogously define $v^*(\theta)$ as the threshold valuation by a buyer such that no purchase is made. Any buyer with a higher value buys; a buyer with lower value sells. Thus, $v^*(\theta)$ is defined by

$$v^* = \theta + \frac{1 - F(v^*)}{f(v^*)} - \int_{v^*-\theta}^{\bar{\epsilon}} (\theta + \epsilon - v^*) dH(\epsilon).$$

The changes in social welfare due to judicial intervention are represented by the following two expressions. The first equation represents the welfare gain from judicial intervention under buyer-designed contracts.

$$\Delta W^{bd} = - \int_{\theta^*(v)}^{\bar{\theta}} \delta(\theta) \left[(v - \theta) H \left(v - \theta - \frac{G(\theta)}{g(\theta)} \right) + \int_{v-\theta-\frac{g}{g}}^{\bar{\epsilon}} \epsilon dH(\epsilon) \right] dG(\theta) dF(v)$$

$$+ \int_{\underline{v}}^{\bar{v}} \int_{\underline{\theta}}^{\theta^*(v)} \left[(v - \theta) \left(H(v - \theta) - H \left(v - \theta - \frac{G(\theta)}{g(\theta)} \right) \right) - \int_{v - \theta - \frac{\epsilon}{g}}^{v - \theta} \epsilon dH(\epsilon) \right] dG(\theta) dF(v). \quad (2.29)$$

$$\begin{aligned} \Delta W^{sd} = & - \int_{\underline{v}}^{v^*(\theta)} \delta(v) \left[(v - \theta) H \left(v - \theta - \frac{1 - F(v)}{f(v)} \right) + \int_{v - \theta - \frac{1 - F(v)}{f(v)}}^{\bar{\epsilon}} \epsilon dH(\epsilon) \right] dF(v) dG(\theta) \\ & + \int_{\underline{\theta}}^{\bar{\theta}} \int_{v^*(\theta)}^{\bar{v}} \left[(v - \theta) \left(H(v - \theta) - H \left(v - \theta - \frac{1 - F(v)}{f(v)} \right) \right) - \int_{v - \theta - \frac{1 - F(v)}{f(v)}}^{v - \theta} \epsilon dH(\epsilon) \right] dF(v) dG(\theta). \end{aligned} \quad (2.30)$$

The former equation represents the welfare gain from judicial intervention under buyer-designed contracts. The latter equation represents the gain under seller-designed contracts. In each equation, the first term in brackets represents the loss from reduced trade; the second term represents the gain from more efficient breach. The central question is under what general conditions are these equations either positive (i.e., judicial intervention is good) or negative (i.e., judicial intervention is bad). Unfortunately, there are no clear general conditions. Rather, the sign of the equations depends fundamentally on the distributions of v , θ , and ϵ . Furthermore, it should be noted that these results depend upon the optimistic assumption that the courts know perfectly the value of loss. If courts make errors, the above loss in welfare may be even greater.¹³

2.5 Extensions and Conclusions

The modeling approach taken in this paper was to assume that the seller may breach with some probability and that the buyer's valuation needed protection from such behavior. Alternatively, we could have chosen an alternative framework where the buyer breaches with some probability and the seller's sunk production costs need to be protected. In this case, we obtain similar results: liquidated damages never exceed the seller's production costs and frequently fail to protect the seller's investment fully. In this sense our explanation regarding the asymmetric treatment of liquidated damage clauses by the courts is robust.

¹³The courts determination of value does not enter the expressions linearly, and so even an unbiased estimate by the court introduces additional nonlinear effects.

This paper has demonstrated that when information of contracting parties is private, liquidated damage clauses serve a dual role of promoting efficient breach and increasing the likelihood of trade. Furthermore, even if the judicial system had perfect information, intervention in the form of prohibiting under-compensatory damages does not necessarily improve social welfare. This may explain why courts have not found it necessary to invalidate under-compensatory damage clauses, but have continued to strike over-compensatory clauses. The former may be the result of a belief that bargaining parties made rational choices, while the latter may be best explained as a belief that excessive damage clauses are symptomatic of contractual failure.

Appendix

Proof of Lemma 1:

Necessity of (2.8) and (2.9):

Incentive compatibility and the definition of $\pi^*(\theta|\theta)$ implies

$$\begin{aligned} \pi^*(\theta|\theta) \geq \pi^*(\hat{\theta}|\hat{\theta}) &+ \delta(\hat{\theta}) \left[\int_{\ell(\hat{\theta})-\theta}^{\ell(\hat{\theta})-\hat{\theta}} \epsilon dH(\epsilon) \right. \\ &\left. + [\ell(\hat{\theta}) - \theta]H(\ell(\hat{\theta}) - \theta) - [\ell(\hat{\theta}) - \hat{\theta}]H(\ell(\hat{\theta}) - \hat{\theta}) \right]. \end{aligned}$$

Integrating by parts and simplifying yields

$$\begin{aligned} \pi^*(\theta|\theta) \geq \pi^*(\hat{\theta}|\theta) &= \pi^*(\hat{\theta}|\hat{\theta}) - \int_{\ell(\hat{\theta})-\theta}^{\ell(\hat{\theta})-\hat{\theta}} \delta(\hat{\theta}) dH(\epsilon), \\ &= \pi^*(\hat{\theta}|\hat{\theta}) - \int_{\hat{\theta}}^{\theta} \delta(\hat{\theta}) H(\ell(\hat{\theta}) - t) dt. \end{aligned}$$

Similarly, $\pi^*(\hat{\theta}|\hat{\theta}) \geq \pi^*(\theta|\theta) + \int_{\hat{\theta}}^{\theta} \delta(\theta) H(\ell(\theta) - t) dt$. Thus, combining the inequalities, we obtain

$$- \int_{\hat{\theta}}^{\theta} \delta(\theta) H(\ell(\theta) - t) dt \geq \pi^*(\theta|\theta) - \pi^*(\hat{\theta}|\hat{\theta}) \geq - \int_{\hat{\theta}}^{\theta} \delta(\hat{\theta}) H(\ell(\hat{\theta}) - t) dt. \quad (2.31)$$

Take $\theta > \hat{\theta}$, without loss of generality, divide (2.31) above by $(\theta - \hat{\theta})$, and take the limit as $\theta \rightarrow \hat{\theta}$. This yields

$$\frac{\pi^*(\theta|\theta)}{d\theta} = -\delta(\theta)H(\ell(\theta) - \theta). \quad (2.32)$$

By (2.31), $\pi^*(\theta)$ is monotonic, and therefore Riemann-integrable, and so we may characterize the seller's profits as

$$\pi^*(\theta|\theta) \equiv \pi^*(\theta) = \pi^*(\bar{\theta}) + \int_{\theta}^{\bar{\theta}} \delta(t)H(\ell(t) - t) dt, \quad (2.33)$$

which is (2.8). Finally, since $\pi^*(\theta)$ is decreasing in θ from (2.33), individual rationality implies that $\pi^*(\theta) \geq 0$. This is (2.9).

Sufficiency of (2.8)-(2.10):

Substituting (2.8) into (2.7) implies $\pi^*(\bar{\theta}) \geq 0$ for all θ ; this is individual rationality. To prove incentive compatibility, note that the definition of $\pi^*(\hat{\theta}|\theta)$ implies

$$\pi^*(\hat{\theta}|\hat{\theta}) = \pi^*(\hat{\theta}|\theta) + \int_{\hat{\theta}}^{\theta} \delta(\hat{\theta}) H(\ell(\hat{\theta}) - t) dt.$$

But (2.8) implies that

$$\pi^*(\bar{\theta}) + \int_{\underline{\theta}}^{\bar{\theta}} \delta(t)H(\ell(t) - t)dt + \int_{\hat{\theta}}^{\theta} \delta(t)H(\ell(t) - t)dt = \pi^*(\hat{\theta}|\theta) + \int_{\hat{\theta}}^{\theta} \delta(\hat{\theta})H(\ell(\hat{\theta}) - t)dt.$$

Applying (2.8) again yields

$$\pi^*(\theta|\theta) = \pi^*(\hat{\theta}|\theta) + \int_{\hat{\theta}}^{\theta} [\delta(\hat{\theta})H(\ell(\hat{\theta}) - t) - \delta(t)H(\ell(t) - t)]dt.$$

By (2.10), the integral above integral is nonnegative which implies $\pi^*(\theta|\theta) \geq \pi^*(\hat{\theta}|\theta)$. Hence, $\{\delta(\theta), p(\theta), \ell(\theta)\}$ is incentive compatible. \square

Proof of Lemma 2:

Necessity of (2.15)-(2.16):

Incentive compatibility and the definition of $\pi^b(v|v)$ implies

$$\pi^b(v|v) \geq \pi^b(\hat{v}|v) = \pi^b(\hat{v}|\hat{v}) + \int_{\underline{\theta}}^{\bar{\theta}} \delta(\hat{v})(v - \hat{v})H(\ell(\hat{v}) - \theta)dG(\theta).$$

Similarly,

$$\pi^b(\hat{v}|\hat{v}) = \pi^b(v|v) + \int_{\underline{\theta}}^{\bar{\theta}} \delta(v)(\hat{v} - v)H(\ell(v) - \theta)dG(\theta).$$

Thus, combining the inequalities, we obtain

$$\int_{\underline{\theta}}^{\bar{\theta}} \delta(v)(v - \hat{v})H(\ell(v) - \theta)dG(\theta) \geq \pi^b(v|v) - \pi^b(\hat{v}|\hat{v}) \geq \int_{\underline{\theta}}^{\bar{\theta}} \delta(\hat{v})(v - \hat{v})H(\ell(\hat{v}) - \theta)dG(\theta). \quad (2.34)$$

Take $v > \hat{v}$, without loss of generality, divide (2.34) above by $(v - \hat{v})$, and take the limit as $v \rightarrow \hat{v}$. This yields

$$\frac{d\pi^b(v|v)}{dv} = \int_{\underline{\theta}}^{\bar{\theta}} \delta(v)H(\ell(v) - \theta)dG(\theta). \quad (2.35)$$

By (2.35), $\pi^b(v)$ is monotonic and therefore Riemann-integrable. Hence, we may characterize the buyer's profits as

$$\pi^b(v|v) \equiv \pi^b(v) = \pi^b(\underline{v}) + \int_{\underline{\theta}}^{\bar{\theta}} \int_{\underline{v}}^{\bar{v}} \delta(t)H(\ell(t) - \theta)dt dG(\theta), \quad (2.36)$$

which is (2.15). Finally, since $\pi^b(v)$ is monotonic, individual rationality implies that $\pi^b(\underline{v}) \geq 0$. This is (2.16).

Sufficiency of (2.15)-(2.17):

Substituting (2.16) into (2.15) implies $\pi^b(\underline{v}) \geq 0$ for all v ; this is individual rationality.

To prove incentive compatibility, note that the definition of $\pi^b(\hat{v}|v)$ implies

$$\pi^b(\hat{v}|\hat{v}) \equiv \pi^b(\hat{v}) = \pi^b(\hat{v}|v) - \int_{\underline{\theta}}^{\bar{\theta}} \delta(\hat{v})(v - \hat{v})H(\ell(\hat{v}) - \theta)dG(\theta).$$

But (21) implies that

$$\begin{aligned} \pi^b(\underline{v}) + \int_{\underline{\theta}}^{\bar{\theta}} \int_{\underline{v}}^{\bar{v}} \delta(t)H(\ell(t) - \theta)dtdG(\theta) + \int_{\underline{\theta}}^{\bar{\theta}} \int_v^0 \delta(t)H(\ell(t) - \theta)dtdG(\theta), \\ = \pi^b(\hat{v}|v) + \int_{\underline{\theta}}^{\bar{\theta}} \delta(\hat{v})(v - \hat{v})H(\ell(\hat{v}) - \theta)dG(\theta). \end{aligned}$$

Applying (2.15) again yields

$$\pi^b(v) = \pi^b(\hat{v}|v) + \int_{\underline{\theta}}^{\bar{\theta}} \int_0^v [\delta(\hat{v})(H(\ell(\hat{v}) - \theta) - \delta(t)H(\ell(t) - \theta))]dtdG(\theta).$$

By (2.17), the double integral above is nonnegative, which implies $\pi^b(v|v) \geq \pi^b(\hat{v}|v)$. Hence, $\{\delta(v), p(v), \ell(v)\}$ is incentive compatible. \square

Proof of Lemma 3:

Following Lemmas 1 and 2, (2.20)-(2.25) are both necessary and sufficient for incentive compatibility and individual rationality. Additionally, (2.20)-(2.23), imply (2.26). To see that (2.24)-(2.26) are sufficient for incentive compatibility and individual rationality, we construct $p(\theta, v)$ such that $\delta(\theta, v)$ and $\ell(\theta, v)$ are IC and IR.

First note that because we have restricted ourselves to piecewise C^1 contracts, $\delta(\theta, v)$ is well defined and $\delta_\theta(\theta, v) = 0$ at all but at a finite number of points. From the envelope theorem, incentive compatibility requires that p must satisfy

$$p_\theta(\theta, v) = [1 - H(\ell(\theta, v) - \theta)]\ell_\theta(\theta, v), \quad (2.37)$$

$$p_v(\theta, v) = [(v - \ell(\theta, v))h(\ell(\theta, v) - \theta) + [1 - H(\ell(\theta, v) - \theta)]\ell_v(\theta, v)], \quad (2.38)$$

whenever $\delta(\theta, v) = 1$. If the constructed price function satisfies these two partial differential equations, we know from Lemmas 1 and 2 that the monotonicity conditions expressed in (2.24)-(2.25) are sufficient for incentive compatibility. One possible construction of p has $p(\theta, v)$ such that $\pi^b(\underline{v}) = 0$. That is,

$$\underline{v}H(\ell(\theta, \underline{v}) - \theta) + \ell(\theta, \underline{v})[1 - H(\ell(\theta, \underline{v}) - \theta)] = p(\theta, \underline{v}).$$

Define,

$$\begin{aligned} p(\theta, v) &\equiv \int_{\underline{\theta}}^{\bar{\theta}} \int_{\underline{v}}^v \{[t - \ell(s, t)]h(\ell(s, t) - s) + [1 - H(\ell(s, t) - s)]\ell_v(s, t)\}dtdG(s) \\ &+ \int_{\underline{\theta}}^{\bar{\theta}} \int_{\underline{v}}^v \{\underline{v}H(\ell(s, \underline{v}) - s) + [1 - H(\ell(s, \underline{v}) - s)]\ell(s, \underline{v})\}dG(s) \end{aligned}$$

$$\begin{aligned}
& - \int_{\underline{v}}^{\bar{v}} \int_{\underline{\theta}}^{\bar{\theta}} [1 - H(\ell(s, t) - s)] \ell_{\theta}(s, t) ds dF(t) \\
& - \int_{\underline{v}}^{\bar{v}} \int_{\underline{\theta}}^{\bar{\theta}} \left\{ [1 - H(\ell(s, t) - s)] \ell_{\theta}(s, t) \frac{G(s)}{g(s)} \right\} dG(s) dF(t).
\end{aligned}$$

The first two expressions represent the expectation over θ of the integral $p_v(\theta, v)$ and the endpoint $p(\theta, \underline{v})$. The second pair of expressions are zero in expectation. It is straightforward to check that (2.37)-(2.38) above are satisfied by this price function, and so the mechanism is incentive compatible. Moreover, (2.26) implies that $\pi^b(\underline{v}) + \pi^s(\bar{\theta}) \geq 0$. And since p was constructed so that $\pi^b(\underline{v} = 0)$, the mechanism is individually rational.

□

References

- AGHION, P. AND BOLTON, P., 1987, "Contracts as a Barrier to Entry," *American Economic Review*, Vol. 77, pp. 388-401.
- DIAMOND, P. AND MASKIN, E., 1979, "An Equilibrium Analysis of Search and Breach of Contract, I: Steady States," *Bell Journal of Economics*, Vol. 10, pp. 282-316.
- FUDENBERG, D. AND TIROLE, J., 1991, *Game Theory*. Forthcoming, M.I.T. Press.
- MASKIN, E. AND TIROLE, J., 1990, "The Principal-Agent Relationship with an Informed Principal, II: Common Values," M.I.T. mimeo.
- MYERSON, R. AND SATTERTHWAITE, M., 1983, "Efficient Mechanisms for Bilateral Trading," *Journal of Economic Theory*, Vol. 29, pp. 265-281.
- MYERSON, R., 1985, "Bayesian Equilibrium and Incomplete Contracts: An Introduction," in Hurwicz, L., et al. (eds.), *Social Goals and Social Organization*. Cambridge University Press.
- POLINSKY, M., 1983, "Risk Sharing Through Breach of Contract Remedies," *Journal of Legal Studies*, Vol. 12, pp. 427-444.
- REA, S., 1984, "Efficiency Implications of Penalties and Liquidated Damages," *Journal of Legal Studies*, Vol. 13, pp. 147-167.
- SHAVELL, S., 1980, "Damage Measures for Breach of Contract," *The Bell Journal of Economics*, Vol. 11, pp. 466-490.
- WILLIAMS, S., 1987, "Efficient Performance in Two Agent Bargaining," *Journal of Economic Theory*, Vol. 41, pp. 154-172.

Chapter 3

Information Expropriation and Moral Hazard in Optimal Auctions

Thrift should be the guiding principle in our government expenditure.

– *Mao Tse-Tung*

The buyer needs a hundred eyes, the seller not one.

– *George Herbert*

3.1 Introduction

In principal-agent environments, “information” rents frequently accrue to agents as a result of their private information. Economists often speak of reducing information rents in these environments through a variety of revelation mechanisms and related devices. One such device is the audit. The decision to audit agents’ reports is based upon the fundamental trade off between the costs incurred from auditing and the gains obtained through reduced information rents. This paper considers a related mechanism for reducing information rents: the transfer of information-inherent “property” from one agent to another. This transfer device, like the audit, involves a similar trade off, but achieves rent reductions by expropriating the agent’s hidden information – transferring part of it

to a competing agent with a lower informational stake in the property. Such an expropriation of information is accomplished by transferring property in which the information is embodied. Providing that an alternative agent has the ability to utilize the asset, such a transfer has the potential for reducing the principal's acquisition costs.

Transferring information is not always possible. For example, in the traditional private-values auction the seller cannot transfer the subjective valuation of an object from one bidder to another. However, in many contexts the transfer of information is a real possibility because such information is embodied in tangible assets. As an illustration of the gains from expropriation of information, this paper develops at length the usefulness of technology transfer in the government procurement context – specifically, in the defense industry.¹ The commitment to transfer, under pre-specified conditions, a defense project from the initial developer to another manufacturer (second source) can be an effective cost-saving strategy for the government.

This mechanism differs from the traditional auction approach which would consider a second source as a bidder with its own cost; in such a case, the optimal auction is straightforward, and may involve handicapping the developer in the auction for production. Here, we assume that while the second source can produce the desired product according to some privately known, independently distributed cost, it can also produce the object using the developer's technology, with the result that its final cost depends less on its own cost structure and more on the developer's private information. This additional option of transferring (or licensing) the developer's technology to the second source may allow the government to reduce its acquisition costs. This paper contributes to auction and contract theory by considering the extent to which the transfer of information-inherent technology reduces information rents in various environments.

This paper begins by examining a simple model in which a tradeoff exists between transfer costs and information rent reductions. In particular, in Sections 2 and 3, I consider the situation of full-commitment power by the buyer in the absence of moral hazard problems by the agents, but where all contracts are constrained to be *ex post*

¹Gansler [1989] provides an interesting overview of many of the more salient issues involved in defense procurement.

individually rational.² The model consists of one buyer (the government) and two sellers (a developer and a second source). The government has three procurement alternatives: choose the developer to produce, choose the second source to produce using her own technology, or choose the second source to produce using technology transferred from the developer. At the contracting stage all parties have symmetric information, and the government commits to specific rules for an auction it will later conduct. The sellers determine their costs and bid accordingly; the rules establish who will produce and how much each seller will receive as a function of the bids.

A rule which optimally utilizes technology transfers induces both the developer and the second source to report their costs truthfully for less information rent. Intuitively, the existence of a second source allows the buyer to compete away some of the information rents via an auction, where the licensing option can be thought of as the addition of an extra seller. Although this additional bidder may have higher costs, it also has less of an informational stake in the transferred technology; if the production costs using the transferred technology are less related to the second source's own costs and more related to the developer's costs, her requisite information rent for truth-telling will be significantly lowered. Consequently, informational gains from technology transfer exist.³

Although a policy of transferring information-laden technology may reduce rents, we might suspect such a policy would have perverse effects upon the developer's initial incentive to invest. In particular, it is plausible that such a policy would reduce the developer's investment in minimizing the project's subsequent cost of production if the buyer can be expected to exercise the option of second-sourcing and partially transfer the developer's efforts. In Section 4, this paper endogenizes the developer's investment decision and derives the optimal auction in the moral hazard environment. There it is assumed that the developer makes an investment which improves the random distribution of the project's production cost. The results indicate that the solution to the moral hazard problem entails a change in the probability of choosing production by the developer as

²Also known as limited liability in the contracts literature. See Sappington [1983].

³The idea of transferring the information-inherent component of one agent to another so as to reduce information rents is not entirely new to the literature. Riordan-Sappington [1989], for example, make use of such transfers in their examination of defense procurement second-sourcing.

a function of the project's reported production cost using the developer's technology. Licensing is implemented more frequently when the developer's announced cost is high than when moral hazard considerations are absent.

The policy ramifications of this paper are twofold. First, this paper demonstrates how a commitment to transferring technology for some bids may reduce expected procurement costs, even when moral hazard is present. Second, this paper provides a caveat for the current empirical practice of evaluating the gains from licensing by comparing the cost of production after a technology transfer with the estimated cost of production by the developer. Such a comparison ignores the *ex ante* gains in reduced information rents which result from the government's commitment to breakout technology for *bad* bids and it ignores the costs of reduced incentives for initial development.

The above framework, although specialized to the procurement setting, is quite general. Following Laffont-Tirole [1988], we may also consider managerial takeover in this framework. We may suppose in period 1 the incumbent managerial team secures profit for its stockholders following a particular profit plan. In period 2, a raider appears who may be employed to takeover the current management team and either institute its own profit strategies, or continue with its predecessors' plans (i.e., plans are transferable).

The analytical approach taken in this paper is that of direct mechanism design for multiple agents under limited liability. In section 2, the model is fully developed and its underlying assumptions are motivated. In section 3, the solution for the optimal contract is developed and Section 4 extends the model by considering moral hazard possibilities by the primary source. Section 5 concludes.

3.2 The Model

We present a model of a risk-neutral buyer with full-commitment ability and two risk-neutral sellers who are subject only to limited-liability. For exposition, we initially consider the problem of a buyer (the government) who must procure an item from two potential sellers (firms). The government desires to procure a single object at the lowest possible cost. It proposes a take-it-or-leave-it contract to both sellers: the primary seller

(firm 1) and the secondary seller (firm 2). The contract commits the buyer to deal with the sellers in a pre-specified manner after the sellers have announced their costs, and must guarantee both sellers non-negative income. Each firm either accepts or rejects the contract. Following their decision, they discover their production costs. The government does not observe costs either *ex ante* or *ex post*.⁴ After learning their costs, firms make announcements (i.e., “bids”) to the government, who chooses which firm will produce; in the case that the second source is chosen, the government additionally chooses whether or not to transfer the developer’s technology. We assume the government’s valuation is sufficiently large that it always chooses to procure the item. Monetary payments are made in accordance with the initial contract.

We may think of the contract that the buyer offers as a commitment to use a specific auction mechanism. In this way, we analyze the problem of choosing the optimal contract as one of optimal auction design. In particular, we will consider truthful revelation mechanisms, using techniques similar to those found in Myerson [1981].

Each firm’s cost, c_i , is independently distributed according to the continuous probability density, $f_i(c_i) > 0$, on a compact set which we take to be $[0, 1]$ without loss of generality. $F_i(c_i)$ is the corresponding cumulative distribution function, and we make the common regularity assumption that $\frac{F_i(c_i)}{f_i(c_i)}$ is nondecreasing in c_i .

The assumption that a firm draws its cost from a distribution can be motivated as the exogenous design of a prototype by defense contractors. Engineers and scientists develop a prototype to meet specific form, fit and function requirements of the government. It is only after the design has been chosen that costs are determined. In this sense, the cost is exogenous. Although we take such draws of cost as exogenous and costless to the firm, we could alternatively assume that the determination of cost entails some fixed amount to build a prototype, prepare a bid, etc. Providing that this amount is small, the government would be willing to pay the known development costs up front to secure the firms’ participation in the auction.

⁴Laffont-Tirole [1986] and McAfee-McMillan [1986] consider contexts where the government can observe costs *ex post*, but is unable to observe the firms’ effort levels. The approach taken here differs from theirs because cost remains unobservable, but similar gains from technology transfers could be realized under alternative models with contractible costs.

The total cost to firm 2 of producing with firm 1's technology, i.e. the total cost of production under licensing, is given by the function $\ell(c_1, c_2)$, which includes the cost of transfer, if any. We will further assume that $\frac{\partial \ell(c_1, c_2)}{\partial c_2} = \ell_2$, a constant; that is, $\ell(c_1, c_2)$ is linear in c_2 . This implies the second source's marginal cost effect on the licensed production is independent of the developer's cost, allowing us to separate the information effects from each other. We also assume that $1 > \ell_2 \geq 0$ and $\frac{\partial \ell(c_1, c_2)}{\partial c_1} < 1$. Consequently, the second source has less informational stake in the transferred technology than its own technology.

The cost distributions and ℓ are common knowledge to the buyer and the sellers. We will often consider a particular situation with linear licensing costs.

$$\ell(c_1, c_2) \equiv \lambda c_1 + (1 - \lambda)c_2 + \gamma, \quad (3.1)$$

for $1 \geq \lambda > 0$. With linear licensing costs, a proportion λ of technology is transferable to firm 2 for a fixed transfer cost γ . In the extreme case of perfectly and costlessly transferable technology we have $\ell(c_1, c_2) = c_1$.

Finally, based upon cost reports, the government chooses from one of three possible production alternatives: (i) primary production; (ii) secondary production; and (iii), technology transfer or licensing (i.e., secondary production with technology transfer). For tractability, we do not include the logical fourth possibility of transferring technology from the secondary firm to the primary firm.

It is important to note that in this framework, seller 2 can be required to produce using seller 1's technology, even when it is inferior to seller 2's own technology. In the defense procurement context, this assumption is plausible. The government requires firm 2 to produce firm 1's design rather than its own. Because designs are easily verifiable, this production decision can be enforced. If this is not feasible because the buyer cannot observe which technology is being used, the optimal contract will must take into account additional constraints in general, but the character of the auction remains similar.

Along with the production decision, the government determines payments to each firm based upon their cost reports. A crucial constraint is that the government must guarantee

non-negative profits for both producers for all possible realizations of cost: No policy can be enforced ex post which would unduly harm a truthful defense company. Here we assume that no firm can be forced to accept a loss. This translates into an ex post profit constraint, below which the government cannot legally force either the developer or the second source. This limited-liability constraint prevents the government from effectively buying the project from the sellers for the expected minimum cost of the production among them.

This assumption is justified for several reasons. First, the assumption approximates a firm that is extremely risk averse beyond a certain level of losses. Given that managers are sensitive to excessive losses, it is plausible that the firms' behavior may be risk neutral over a moderate range but risk averse for dramatic losses. Additionally, from a purely descriptive perspective it is doubtful that the government could force a defense company to continue production when it suffers excessively large losses. Boards of Contracts Appeal (BCAs), the neutral tribunals which have jurisdiction over government contract disputes, frequently grant equitable adjustments to contracts which impose excessive sacrifices upon firms. To this extent, a limit exists to the losses which a contractor can be forced to bear.

We do not allow the primary agent's payoffs to depend upon any ex post discoveries made by the secondary agent after a transfer. If the buyer could do this, the first best solution would be approximated by employing the secondary agent with an arbitrarily small probability to check the truth-telling of the primary agent, and then punishing the primary agent sufficiently hard whenever untruthful reports occur. This paper considers the more subtle issue involved when payments cannot be conditioned on an ex post report of another agent. Such a restriction appears realistic in the defense procurement context; otherwise we would have to allow a time delay (perhaps years) between the auction and the agent's action (e.g., defense production) before enough evidence could be marshaled to levy a punishment against the primary agent.

Although the model we consider has only one production stage, this is without loss of generality. Consider a model in which the developer is the only potential producer at the first stage, so the government must buy from developer at the highest plausible cost. No

information is revealed to the buyer. In stage 2 the second source appears as competition, and the timing is as before. Thus, the model we are examining is readily extendible to a two-stage production model. We can cast defense procurement as a process where a developer initially produces the object, and later a second source appears to compete. The government has the option of continuing to buy from the developer, letting the second source produce its own version of the project, or transferring the developer's technology to the second source to produce.

3.3 The Optimal Contract

In this model, the buyer commits to deal with the sellers in a predetermined manner after learning their reported costs. Using these reports, the buyer determines who produces the object, whether technology is transferred, and how much each firm shall be paid. The Revelation Principle states that without loss of generality, we may restrict our attention to direct revelation mechanisms. The class of mechanisms which we will consider is given by $\mathcal{M} = \left\{ \{p_i(c_1, c_2)\}_{i=0}^2, \{t_j(c_1, c_2)\}_{j=1}^2 \right\}$, where, for given reported costs, p_i is the probability that production alternative i is chosen by the buyer, and t_j is the transfer to firm j . The production alternatives, $i = 0, 1, 2$, correspond to licensed production, firm 1 (developer) production, and firm 2 (second source with own technology) production, respectively.

3.3.1 The First Best

Before examining the optimal contract under limited liability and asymmetric information, we note the properties of the full-information contract. Under the full information contract

(i) the most efficient form of production is chosen:

$$p_0(c_1, c_2) = \begin{cases} 1 & \text{if } \ell(c_1, c_2) \leq \min\{c_1, c_2\}, \\ 0 & \text{otherwise,} \end{cases}$$

$$p_1(c_1, c_2) = \begin{cases} 1 & \text{if } c_1 < \min\{c_2, \ell(c_1, c_2)\}, \\ 0 & \text{otherwise,} \end{cases}$$

$$p_2(c_1, c_2) = \begin{cases} 1 & \text{if } c_2 < \min\{c_1, \ell(c_1, c_2)\}, \\ 0 & \text{otherwise;} \end{cases}$$

(ii) the buyer pays the producer realized cost:

$$t_1(c_1, c_2) = \begin{cases} c_1 & \text{if } p_1(c_1, c_2) = 1, \\ 0 & \text{otherwise,} \end{cases}$$

$$t_2(c_1, c_2) = \begin{cases} c_2 & \text{if } p_2(c_1, c_2) = 1, \\ \ell(c_1, c_2) & \text{if } p_0(c_1, c_2) = 1, \\ 0 & \text{otherwise;} \end{cases}$$

(iii) the firms make zero profit.

Because there is full information, the limited-liability constraint is not binding, as zero profits may be guaranteed for all outcomes. The firms will be willing to accept the above contract, and the buyer obtains the object at minimum (in this case, actual) cost. Any contract yielding a lower expected price must necessarily violate individual rationality. Note that if $\ell(c_1, c_2) > \min\{c_1, c_2\}$ for all c_1, c_2 , then licensing is never optimal under full-information. We will see below that even when licensing would never be optimal under full-information, licensing may be a desirable strategy by the buyer in environments of asymmetric information.

3.3.2 Asymmetric Information and Limited Liability

The ability to commit will be important when sellers have private information about costs because, *ex post*, the government often will not desire to use the possibly inefficient licensed second source. By committing *ex ante* to license for some given cost reports, the government may reduce the expected costs of purchase.

Under the assumption that the other firm is truthful, payoffs to each firm as a function of reported and true costs are

$$\pi_1(\hat{c}_1, c_2 | c_1) \equiv t_1(\hat{c}_1, c_2) - p(\hat{c}_1, c_2)c_1,$$

$$\pi_2(c_1, \hat{c}_2 | c_2) \equiv t_2(c_1, \hat{c}_2) - p_2(c_1, \hat{c}_2)c_2 - p_0(c_1, \hat{c}_2)\ell(c_1, c_2),$$

where $\hat{\cdot}$ denotes the reported type. Because neither firm knows the others cost when it

must make its report, it is useful to consider the expected payoffs for each firm:

$$\pi(\hat{c}_1|c_1) = \int_0^1 \{t_1(\hat{c}_1, c_2) - p_1(\hat{c}_1, c_2)c_1\} dF_2(c_2), \quad (3.2)$$

$$\pi(\hat{c}_2|c_2) = \int_0^1 \{t_2(c_1, \hat{c}_2) - p_2(c_1, \hat{c}_2)c_2 - p_0(c_1, \hat{c}_2)\ell(c_1, c_2)\} dF_1(c_1), \quad (3.3)$$

The mechanism-design problem facing the buyer is given below as program P1:

$$\min_{\mathcal{M}} \int_0^1 \int_0^1 \{t_1(c_1, c_2) + t_2(c_1, c_2)\} dF_1(c_1)dF_2(c_2) \quad (3.4)$$

subject to

$$\pi_j(c_j|c_j) \geq \pi_j(\hat{c}_j|c_j), \forall c_j, \hat{c}_j, \quad (3.5)$$

$$\pi_j(c_1, c_2|c_j) \geq 0, \forall c_1, c_2. \quad (3.6)$$

The objective function is the expected value of the payments paid by the government for the procurement. This is minimized subject to constraints (3.5) and (3.6). Constraints in (3.5) ensure Bayesian truth-telling. Constraints in (3.6) represent the limited-liability constraints for all states of nature; note that this is not an expectation over payments, but actual payment.

Following Mirrlees [1971], Myerson [1981], et al., we simplify the truth-telling and limited-liability constraints, and incorporate them into the objective function to ascertain the nature of the optimal auction to obtain our first result.

Proposition 1 *The set of $\{p_i(c_1, c_2)\}_{i=0}^2$ which solve P1 is the same as that which solves program P2 below using point-wise minimization over the space of probability distributions on the production alternatives:*

$$\min_{p_0, p_1, p_2} \left\{ p_1 \left[c_1 + \frac{F_1(c_1)}{f_1(c_1)} \right] + p_2 \left[c_2 + \frac{F_2(c_2)}{f_2(c_2)} \right] + p_0 \left[\ell(c_1, c_2) + \ell_2 \frac{F_2(c_2)}{f_2(c_2)} \right] \right\}. \quad (3.7)$$

Proof: see Appendix.

Note that when $\ell(c_1, c_2) = c_2$, the Proposition reduces to the standard auction result

which may involve handicapping if the cost distributions differ. To see the mechanics of this solution to the optimal auction, define the following variables as the virtual costs of each production alternative:

$$\begin{aligned} J_1(c_1, c_2) &\equiv c_1 + \frac{F_1(c_1)}{f_1(c_1)}, \\ J_2(c_1, c_2) &\equiv c_2 + \frac{F_2(c_2)}{f_2(c_2)}, \\ J_0(c_1, c_2) &\equiv \ell(c_1, c_2) + \ell_2 \frac{F_2(c_2)}{f_2(c_2)}. \end{aligned}$$

Thus, the solution to P2 amounts to selecting the alternative with the minimum virtual-cost. It will also be useful for a graphical analysis to define the following state-space partition over the set of all possible realizations of cost, where Ω^i is the set of (c_1, c_2) such that alternative i has the lowest virtual cost. That is,

$$\begin{aligned} (c_1, c_2) \in \Omega^0 &\Leftrightarrow J_0(c_1, c_2) < \min\{J_1(c_1, c_2), J_2(c_1, c_2)\}, \\ (c_1, c_2) \in \Omega^1 &\Leftrightarrow J_1(c_1, c_2) \leq \min\{J_0(c_1, c_2), J_2(c_1, c_2)\}, \\ (c_1, c_2) \in \Omega^2 &\Leftrightarrow J_0(c_1, c_2) \leq \min\{J_0(c_1, c_2), J_1(c_1, c_2)\} \\ &\quad \text{and } J_1(c_1, c_2) \neq J_2(c_1, c_2). \end{aligned}$$

The following Corollary flows directly from the definitions and the optimization of P2 in Proposition 1.

Corollary 1 *The optimal auction consists of setting $p_i(c_1, c_2) = 1$ iff $(c_1, c_2) \in \Omega^i$.*

The sets Ω^0, Ω^1 , and Ω^2 , represent cost realizations where licensing, developer production, and second-source production are chosen, respectively. Note that it is never strongly optimal to randomize between alternatives. The payments which implement the choices in P2 are determined using standard techniques.

Proposition 2 *Optimal payments which correspond to the solution of P2 are given by*

$$t_1(c_1, c_2) = \int_{c_1}^1 p_1(c_1, c_2) dc_1 + p_1(c_1, c_2)c_1, \quad (3.8)$$

$$\begin{aligned}
t_2(c_1, c_2) = & \int_{c_2}^1 [p_0(c_1, c_2)\ell_2 + p_2(c_1, c_2)] dc_2 \\
& + p_2(c_1, c_2)c_2 + p_0(c_1, c_2)\ell(c_1, c_2).
\end{aligned} \tag{3.9}$$

Proof: See Appendix.

Note that in all but the worst states, the above payment scheme pays positive rents to the firm chosen to produce, while the other firm receives nothing.

3.3.3 The Value of Technology Transfers

The commitment to use technology transfers under some cost realizations reduces ex ante information rents by relaxing firm 2's incentive compatibility constraints. Firm 2 can "less easily" say that it has high costs, because the buyer can always transfer firm 1's technology for it to produce.

To understand the intuition behind the optimal auction, consider the following polar case: $\ell(c_1, c_2) = c_1$. That is, firm 1's technology is completely and costlessly transferred under licensing to firm 2. For symmetry in this case, also assume technology can be transferred from firm 2 to firm 1, completely and costlessly. Now a buyer may offer the following contract to extract fully the rent: If $c_1 \leq c_2$, transfer firm 1's technology to firm 2 and have firm 2 produce the project using firm 1's technology for payment c_1 ; if $c_1 > c_2$, vice versa. Under this scheme, neither firm has an incentive to lie and the buyer completely extracts the information rents. Moreover, this scheme does not require firms to know each others costs at the time of bidding.

Returning to our one-way technology transfer environment, transfers of technology under the optimal contract are ex ante optimal whenever $(c_1, c_2) \in \Omega^0$. An interesting question regards the determination of this region. Essentially the buyer trades off the costs of inefficient licensing against the gain in reduced information rents. This is easily seen in the following Proposition.

Proposition 3 *The ex ante expected gain to the buyer from a policy of optimal licensing*

is given by

$$\int_{\Omega_1^0} \left(\frac{F_1(c_1)}{f_1(c_1)} - \ell_2 \frac{F_2(c_2)}{f_2(c_2)} \right) dF_1(c_1)dF_2(c_2) + \int_{\Omega_2^0} (1 - \ell_2) \frac{F_2(c_2)}{f_2(c_2)} dF_1(c_1)dF_2(c_2) \\ - \int_{\Omega_1^0} (\ell(c_1, c_2) - c_1) dF_1(c_1)dF_2(c_2) - \int_{\Omega_2^0} (\ell(c_1, c_2) - c_2) dF_1(c_1)dF_2(c_2),$$

where Ω_i^0 is the licensing region where alternative i would have been chosen if licensing were not available; i.e., where $J_0 < J_i \leq J_{-i}$.

Proof: The result follows from noting that the gain from licensing is the expected reduction in virtual cost from licensing over a standard optimal auction without technology transfer. Since chosen virtual costs are only changed over Ω^0 , we take expectations over this space. The expression immediately follows. \square

The Proposition identifies two effects. The first two terms represent the gain to the buyer from information rent reductions. The last two terms represent the cost inefficiencies to the buyer from deciding on an inefficient production technique. The optimal contract can be reformulated as one in which Ω^0 maximizes the sum of the terms. If no Ω^0 exists such that the sum is positive, the optimal contract does not entail licensing for any realization of costs. This suggests a Corollary.

Corollary 2 *If $\ell(c_1, c_2) = \lambda c_1 + (1 - \lambda)c_2 + \gamma$, $\lambda \in (0, 1)$, and the sellers' cost distributions are symmetric on $[\underline{c}, \bar{c}]$, then an optimal auction will transfer technology with positive probability if*

$$\lambda > \gamma f(\bar{c}).$$

If costs are distributed uniformly on $[\underline{c}, \bar{c}]$, then the optimal contract will utilize transfers if $\lambda(\bar{c} - \underline{c}) > \gamma$.

This result is in contrast to the result in Riordan-Sappington [1989] who find in their model without limited-liability constraints and without commitment that second sourcing is rarely optimal. Because Riordan-Sappington do not assume limited-liability, the firms

compete away expected information rents at the initial symmetric information stage, so there is no information-rent problem. The gains from technology transfer in their model do not derive from reductions in information rents, but from production enhancement: the government introduces less distortion in its decision of whether to produce at all if a second source exists as an alternative. This latter effect is absent in the present model because we have assumed for tractability that the government always procures the object – otherwise, we would find an additional positive term in Proposition 3, providing another gain to technology transfers.

3.3.4 An Example

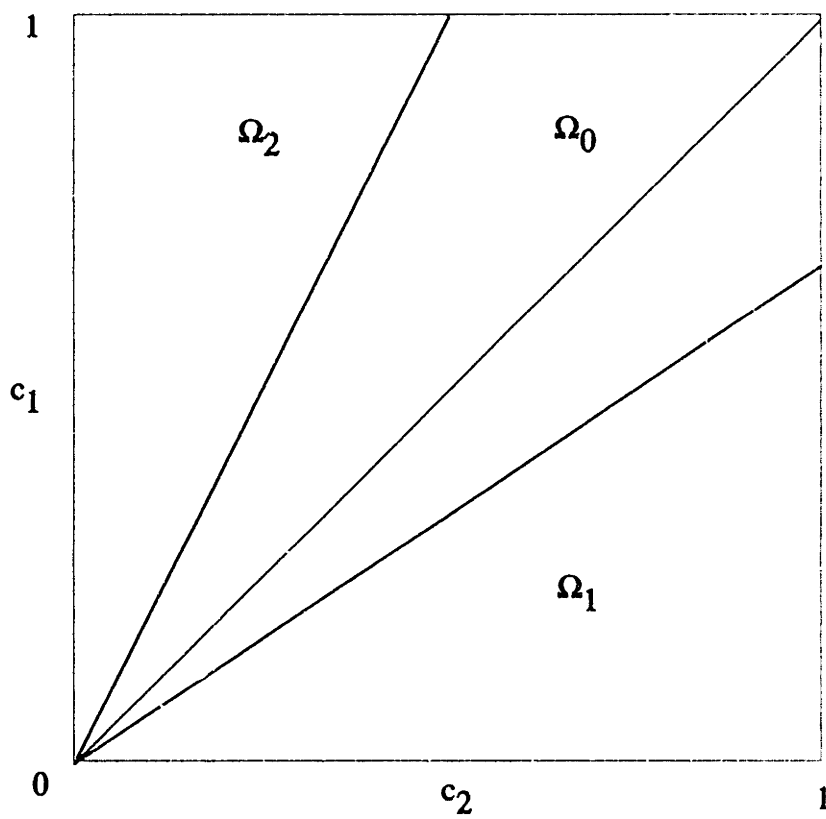
Consider the following linear cost model with uniform distributions on $[0,1]$. That is, let $F_i(c_i) = c_i, i = 1, 2$, and let $\ell(c_1, c_2) = \lambda c_1 + (1 - \lambda)c_2 + \gamma$. Thus the virtual costs are given by

$$\begin{aligned} J_1(c_1, c_2) &= 2c_1, \\ J_2(c_1, c_2) &= 2c_2, \\ J_0(c_1, c_2) &= \lambda c_1 + 2(1 - \lambda)c_2 + \gamma. \end{aligned}$$

For the initial case, we make the further simplifying assumptions that $\lambda = \frac{1}{2}$ and $\gamma = 0$. The optimal partition over $[0,1]$ is graphed in Figure 1 as the projection of the minimum virtual cost onto the cost space.

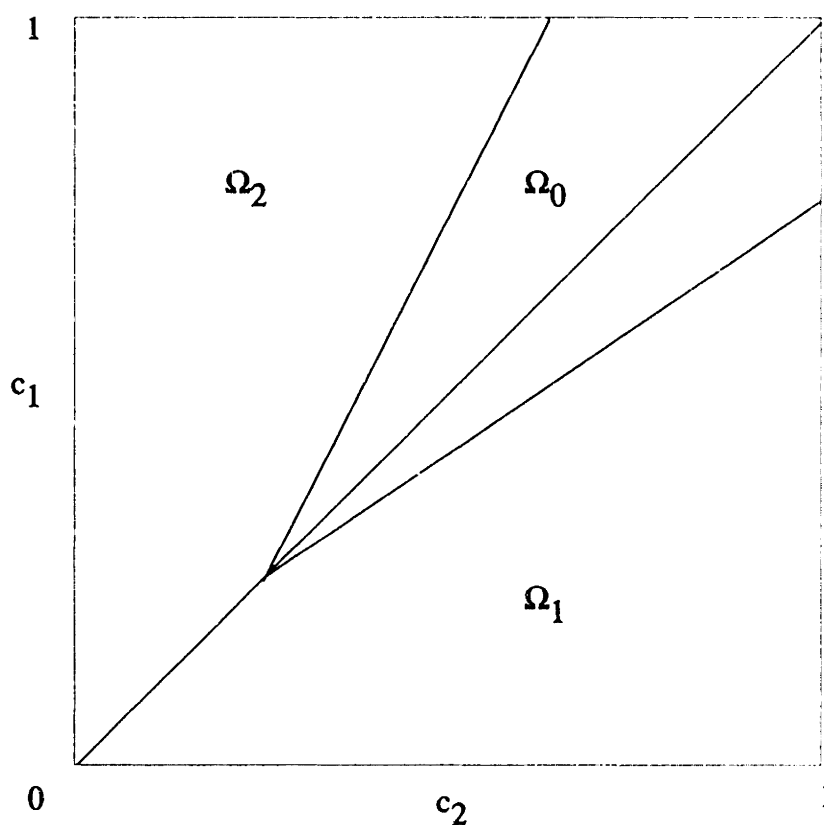
The diagram indicates that when cost reports are relatively close, licensing is chosen. Intuitively, if the cost reports are relatively close, the licensing cost does not differ significantly from either the developer or second source production, so there is little cost inefficiency from licensing. If firm 2 has a relatively low cost, it is expensive for the buyer to make the second source use the inefficient licensed technology rather than its own. Similarly, if firm 1 has a relatively low cost, it is productively inefficient to license technology to firm 2, since firm 1 is a superior producer. As costs become close, the losses in production inefficiencies shrink to zero and are offset by the gains from reduced information rents.

FIGURE 1. OPTIMAL AUCTION: $\lambda = \frac{1}{2}, \gamma = 0$.



We would expect the introduction of a fixed cost for transfer to increase the productive inefficiencies associated with licensing, and consequently the state space associated with licensing to shrink. To see the effect of a transfer cost, consider fixed licensing costs of $\gamma = \frac{1}{8}$.

FIGURE 2. OPTIMAL AUCTION: $\lambda = \frac{1}{2}, \gamma = \frac{1}{8}$.



The licensing region has decreased substantially. As Corollary 2 predicts, if $\ell(c_1, c_2) = \lambda c_1 + (1 - \lambda)c_2 + \gamma$ and costs are symmetrically and uniformly distributed on $[0,1]$, then there is no gain to licensing when $\gamma \geq \lambda$. As γ increases to $\lambda = \frac{1}{2}$, the optimal licensing area shrinks to zero. More generally, Proposition 3 indicates that an increase in $\ell(c_1, c_2)$ (holding ℓ_2 , c_1 , and c_2 fixed) will reduce the probability of licensing and, if the increase is sufficiently large, will eliminate its use altogether; mathematically, the costs of licensing (the latter terms in Proposition 3) increase while the benefits (the former terms) remain

unchanged.

As indicated, the above analysis has clear applications to the problems of procurement and the question of whether the government should second source a project by transferring the technology. The analysis suggests that technology transfer might reduce the government bill *ex ante* if appropriately administered. It should be recognized, however, that to the extent that firm (e.g., the developer) may directly affect the transferability of its technology by expending effort in idiosyncratic design, we might expect that firms may engage in wasteful investment to increase transfer costs, and thus licensing will become a less viable alternative. If this is a possibility, a policy of technology transfer must be carefully evaluated.

3.4 Moral Hazard

We naturally expect that in some situations where the initial agent (the developer) must make unobservable investments in reducing the marginal cost, c_1 , of the final product, a policy of expropriating information via technology transfer would induce significant moral hazard. If the buyer can freely transfer the design to a second source to produce, the primary agent may have less incentive to reduce the marginal cost of production. Consequently, an examination of the moral hazard dimension seems particularly relevant.

3.4.1 The Problem of Moral Hazard

This section extends the previous analysis by incorporating moral hazard on the part of the primary agent. We model this extension by assuming that the primary agent (the developer) may make cost-reducing investments. Here we consider only two production choices for simplicity: The buyer may decide to purchase from either the developer or the licensed second source (using the developer's technology). The question we ask is whether the buyer will find it optimal to favor the developer for cost-reducing investments in the award of the production contract, and if so, how?

As before, the approach we take is one of full-commitment by the buyer and limited-liability constraints for the sellers. Initially, the buyer proposes a contract to the two

sellers which is accepted if it guarantees each nonnegative profit in all states of nature. Following the offer the developer chooses cost-reducing investment, e . This investment stochastically shifts (in a first-order sense) the distribution of the developer's production costs, c_1 , and thereby improves the licensed cost of production as well. After investments have been made, costs of production are drawn by each firm from known distributions, with each firm's actual cost being observed only by that individual firm. The sellers then report their costs to the government. The government follows the agreed-upon contract and awards the production decision and payments conditional on the project's valuation and the cost announcements.

The resulting optimal contract is found to be a variation of the classical optimal auction design which awards production to the most favorable virtual type. Under moral hazard, we find that the developer's virtual type has an additional term which decreases in production cost in a manner closely akin to the sharing rule in Holmström [1979]. This suggests that in the stochastic cost-investment model, we would expect a discriminating auction to be utilized which may additionally favor the developer depending upon the resulting cost realizations.

3.4.2 The Model with Moral Hazard

The cost to firm 2 of producing with firm 1's technology is as before. The cumulative distribution function for the developer's cost is now given by $F_1(c_1|e)$, and it is assumed that $\frac{\partial F_1(c_1|e)}{\partial e} \geq 0$; that is, effort leads to a first-order stochastic improvement in the distribution on costs. For tractability, we will assume $F_1(c_1|e)$ satisfies the Concave Distribution Function Condition (CDFC): $F_{1,ee} = \frac{\partial^2 F_1(c_1|e)}{\partial e^2} \leq 0$ for all e, c_1 . This condition assures us that the first-order approach to the principal-agent problem is valid.⁵

The cost to the developer for value-enhancing effort is given by $\psi(e)$, where $\psi(e)$ is

⁵A Monotone Likelihood Ratio Condition (MLRP) is usually required in pure moral hazard settings in addition to CDFC in order to assure that the agent's payoffs are monotonic in outcome. See Grossman-Hart [1983], Rogerson [1985] for proofs of this proposition. [We use *concavity* in distributions rather than *convexity* as in Grossman-Hart, because higher costs are considered undesirable in our model.] With adverse selection, incentive compatibility requires that $\pi(c_1)$ be nonincreasing, and so we do not need a MLRP condition for sufficiency in the first-order approach. We may, however, have to solve the buyer's program subject to monotonicity of π in costs.

increasing, strictly convex, $\psi'''(e) > 0$, $\psi(0) = \psi'(0) = \psi''(0) = 0$, and $\psi(1) = \infty$.

Finally, based upon cost reports the government chooses from one of two possible production alternatives: (i) licensed production; or (ii) developer production. Along with the production decision, the government determines payments to each firm based upon their cost reports. Again the crucial constraint is that the government must guarantee non-negative profits for both producers for all possible realizations of cost.

3.4.3 The Optimal Contract under Moral Hazard

The class of mechanisms considered is given by $\mathcal{M}' = \left\{ \{p_i(c_1, c_2)\}_{i=0}^1, \{t_j(c_1, c_2)\}_{j=1}^2 \right\}$, analogous to before. The production alternatives, $i = 0, 1$, correspond to licensed production and developer production, respectively.

The Choice of Investment. Consider first the investment decision. Given the assumptions regarding the distribution of costs, the developer's choice of effort solves

$$\max_{e \in [0,1]} \int_0^1 \int_0^1 \pi_1(c_1, c_2) dF(c_2) dF_1(c_1|e) - \psi(e).$$

We can more simply characterize the solution to this program in the following Result.

Result 1 *The necessary and sufficient condition of the solution to the agent's effort decision is*

$$\int_0^1 \int_0^1 \left\{ p_1(c_1, c_2) \frac{F_{1,e}(c_1|e)}{f_1(c_1|e)} \right\} dF_1(c_1|e) dF_2(c_2) = \psi'(e), \quad (3.10)$$

where e subscripts denote partial derivatives with respect to e .

Proof: The first-order condition for the solution is

$$\int_0^1 \int_0^1 \pi_1(c_1, c_2) f_{1,e}(c_1|e) dc_1 dF_2(c_2) - \psi'(e) = 0. \quad (3.11)$$

A sufficient condition for a maximum is that

$$\int_0^1 \int_0^1 \pi_1(c_1, c_2) f_{1,ee}(c_1, e) dc_1 dF_2(c_2) - \psi''(e) < 0,$$

for all e . Integrating this expression by parts, and noting that Lemma 1 from the Appendix (used in the proof of Proposition 1) implies $\frac{\partial \pi_1(c_1, c_2)}{\partial c_1} = -p_1(c_1, c_2)$, yields an equivalent condition,

$$\int_0^1 [\pi_1(c_1, c_2) F_{1,ee}(c_1|e)]_0^1 dF_2(c_2) + \int_0^1 \int_0^1 p_1(c_1, c_2) F_{1,ee}(c_1|e) dc_1 dF_2(c_2) - \psi''(e) < 0.$$

CDFC and the strict convexity of $\psi(e)$ assures us that the second-order condition for a maximum holds, thus (3.11) is both necessary and sufficient. Integrating by parts yields:

$$\int_0^1 \int_0^1 -\frac{\partial \pi_1(c_1, c_2)}{\partial c_1} \frac{F_{1,e}(c_1|e)}{f_1(c_1|e)} dF_1(c_1|e) dF_2(c_2) - \psi'(e) = 0.$$

Substituting in the incentive compatibility condition from Lemma 1 in the Appendix and we get the desired result. \square

General Solution to the Contracting Problem. Having characterized the effort chosen by the developer for a given contract we compute the buyer's optimal contract in the presence of moral hazard. To do so we simply append to the buyer's problem the additional condition from in Result 1 to endogenize the investment decision. Call this program P3, and let μ represent the Lagrange multiplier associated with the investment constraint. Proposition 4 below provides the equivalence of P3 with a point-wise minimization problem. Proposition 5 further characterizes the optimal contract.

Proposition 4 *Assume that $\frac{F_{1,e}(c_1|e)}{f_1(c_1|e)}$ is nondecreasing in c_1 , $\frac{F_1(c_1|e)}{f_1(c_1|e)}$ is nonincreasing in e , and $F_{1,ee}(c_1|0) = \frac{\partial}{\partial e} \left(\frac{F_1(c_1|0)}{f_1(c_1|0)} \right) = 0$. The set of $\{p_i(c_1, c_2)\}_{i=0}^1$ which solves P3 is the same as that which solves*

$$\min_{p_i} \left\{ p_1 \left[c_1 + \frac{F_1(c_1|\hat{e})}{f_1(c_1|\hat{e})} + \mu \frac{F_{1,e}(c_1|\hat{e})}{f_1(c_1|\hat{e})} \right] + p_0 \left[\ell(c_1, c_2) + \ell_2 \frac{F_2(c_2)}{f_2(c_2)} \right] \right\}$$

using point-wise maximization, where \hat{e} is the buyer's expectation of firm 1's effort (which is correct in equilibrium) and $\mu > 0$. The level of effort, \hat{e} , induced by the buyer satisfies (3.10), and μ and \hat{e} jointly satisfy

$$\begin{aligned}
& \mu \left\{ \int_0^1 \int_0^1 p_1(c_1, c_2) \left(\frac{F_{1,ee}(c_1|\hat{e})}{f_1(c_1|\hat{e})} \right) f_1(c_1|\hat{e}) dc_1 dF_2(c_2) - \psi''(\hat{e}) \right\} \\
& \quad + \int_0^1 \int_0^1 p_1(c_1, c_2) \frac{\partial}{\partial e} \left(\frac{F_1(c_1|\hat{e})}{f_1(c_1|\hat{e})} \right) f_1(c_1|\hat{e}) dF_2(c_2) \\
& = \psi'(\hat{e}) - \int_0^1 \int_0^1 \left[\sum_{i=0}^1 p_i(c_1, c_2) \tilde{J}_i(c_1, c_2) \right] f_{1,e}(c_1|\hat{e}) dc_1 dF_2(c_2), \quad (3.12)
\end{aligned}$$

where $\tilde{J}_1(c_1, c_2) \equiv c_1 + \frac{F_1(c_1|e)}{f_1(c_1|e)} + \mu \frac{F_{1,e}(c_1|e)}{f_1(c_1|e)}$ and $\tilde{J}_0(c_1, c_2) \equiv \ell(c_1, c_2) + \ell_2 \frac{F_2(c_2)}{f_2(c_2)}$.

Proof: See Appendix.

The assumptions for the Proposition regarding the monotonicity of $\frac{F_{1,e}}{f_1}$ in c_1 and $\frac{F_1}{f_1}$ in e are satisfied if e has more effect on reducing higher cost levels and the developer cannot increase information rents (i.e., the inverse hazard rate) by increasing investment. Following the above Proposition, transfers are constructed to satisfy the limited liability constraint as below.

Proposition 5 *A set of transfers corresponding to the optimal contract are given by*

$$\begin{aligned}
t_1(c_1, c_2) &= \int_{c_1}^1 p_1(c_1, c_2) dF_2(c_2) + p_1(c_1, c_2) c_1 + \psi(\hat{e}), \\
t_2(c_1, c_2) &= \int_{c_2}^1 p_2(c_1, c_2) \ell_2 dF_1(c_1|\hat{e}) + p_2(c_1, c_2) \ell(c_1, c_2).
\end{aligned}$$

Proof: Follows from Proposition 4. □

The solution to the principal's problem has the same nature as the optimal auction without moral hazard, except that the state-space partition over firm production has been changed in an important way – it now depends more intricately upon the realization of the developer's cost. Consider the developer's virtual cost to the buyer:

$$\tilde{J}_1(c_1, c_2) = c_1 + \frac{F_1(c_1|e)}{f_1(c_1|e)} + \mu \frac{F_{1,e}(c_1|e)}{f_1(c_1|e)}.$$

There is an additional term in the virtual cost that was not present before which is very

similar to the optimal sharing rule in Holmström [1979]. This new term represents an additional reward for cost reduction that the developer receives through departures from bidding parity in the auction for production. This additional term serves to increase the sensitivity of the developer's virtual cost by increasing the marginal effect of a reduction in c_1 and thereby increasing $p_1(c_1)$. Furthermore, we know that the moral hazard term $F_{1,c}/f_1$ must be nonpositive, indicating that the developer is favored in the auction. Of course, the buyer realizes the developer did not shirk under the optimal scheme, but nevertheless the buyer must commit to "over" reward the developer for low costs if she wishes to maximize surplus from an ex ante point of view.

The additional term in the virtual cost of the developer reflects the interdependence of the moral hazard and adverse selection problems in this model. Rewards for low costs are accomplished by appropriately tilting the incentive scheme. Unlike Holmström, in our case rewards are made by changing the probability of winning the auction rather than through lump-sum payments.

3.5 Conclusions

The immediate implications of the above analysis suggest that a committed policy of technology transfer is a useful device for reducing information rents. Moreover, the role of transfer is more than simple monitoring. No information of the primary agent needs to be known by a secondary agent for such a transfer to yield benefits for a principal. The policy implications suggest it may be optimal to switch to a possibly inefficient agent in order to reduce ex ante rents. Finally, any empirical test of such behavior must carefully evaluate the strategic effects upon the agents, or it is possible that an optimal transfer policy will appear wasteful.

In the context of managerial incentives, Scharfstein [1988] examines the disciplinary role of a corporate raider who is informed of the firm's true value, and finds such an informed raider both induces incumbent managers to work harder and reduces their information rents. His model is closely analogous to this paper in that the firm value (known by the incumbent managers) transfers completely to the raider if there is a takeover. This

paper demonstrates that while a raider is more effective in reducing information rents if she knows the incumbent's information, there is nonetheless a positive role for uninformed raiders in reducing information rents. There is no requirement that the alternative agent have any *ex ante* knowledge of the primary agent's cost realization for information rents to be reduced.

When investment concerns are present, the optimal auction will favor the developer depending upon cost realizations. As in the pure adverse selection environment, empirical tests of supplier switching under an optimal regime with moral hazard concerns may also yield negative results. One must be very careful to evaluate a policy of source switching from an *ex ante* perspective where its true usefulness may be better understood.

Related to this work is that of Riordan-Sappington [1989]. They consider a model of effort-enhanced value, in their no-commitment, unlimited-liability environment. Because the buyer cannot commit, the developer can expect the buyer to behave opportunistically after investment is sunk. Under this framework, the inability to commit not to use a second source leads to inefficient investment in most plausible cases. If commitment were possible, the government could promise to purchase the product at a price equal to its valuation and let the potential sellers bid away the expected information rents *ex ante* in the competition for the development contract at the symmetric information stage. In this paper, the limited-liability constraint implies any gain from information rent reduction is a direct gain to the buyer. The tradeoffs involved are very different.

Laffont-Tirole [1988] also consider a dynamic adverse selection-moral hazard framework. They find that if investment is completely transferable from the developer to the second source, the buyer would do best to commit to favor the developer at the competition for determining the producer. The results are similar in that bidding parity is disposed of to provide incentives for value-enhancing, transferable investment.

Appendix

Proof of Propositions 1 and 2: The proof of Proposition 1 proceeds with three Lemmas. Lemma 1 establishes necessary and sufficient conditions for truth-telling (3.5) and interim individual rationality (IIR), a weaker constraint than (3.6); the IIR constraint is given by:

$$\pi_j(c_j|c_j) \geq 0, \forall c_j, j = 1, 2.$$

Lemma 2 establishes that the modified program of minimizing (3.4) over these new conditions is equivalent to solving P2 point-wise. Finally, Lemma 3 shows that a particular solution to the modified program is "equivalent" to the solution of P1.

For notational convenience, we will sometimes denote a function which has had expectations taken over one argument, as a function of only the single remaining argument. E.g., $p_1(c_1) \equiv \int_0^1 p_1(c_1, c_2) dF_2(c_2)$.

Lemma 1 *Incentive compatibility (IC) and interim individual rationality (IIR) hold if and only if*

$$\pi_1(c_1|c_1) = \pi_1(1|1) + \int_{c_1}^1 p_1(c_1) dc_1, \quad (3.13)$$

$$\pi_2(c_2|c_2) = \pi_2(1|1) + \int_{c_2}^1 [p_2(c_2) + p_0(c_2)\ell_2] dc_2, \quad (3.14)$$

$$p_1(c_1) \geq p_1(c'_1), \forall c'_1 > c_1, \quad (3.15)$$

$$p_2(c_2) + p_0(c_2)\ell_2 \geq p_2(c'_2) + p_0(c'_2)\ell_2, \forall c'_2 > c_2, \quad (3.16)$$

$$\pi_1(1|1) \geq 0, \quad (3.17)$$

$$\pi_2(1|1) \geq 0. \quad (3.18)$$

Proof:

Necessity.

Consider firm 1. IC and the definition of $\pi_1(\hat{c}_1|c_1)$ implies

$$\pi_1(c_1|c_1) \geq \pi_1(\hat{c}_1|c_1) = \pi_1(\hat{c}_1|\hat{c}_1) - p_1(\hat{c}_1)(c_1 - \hat{c}_1).$$

Rearranging and reversing the roles of c_1 and \hat{c}_1 yields

$$-p_1(c_1)(c_1 - \hat{c}_1) \geq \pi_1(c_1|c_1) - \pi_1(\hat{c}_1|\hat{c}_1) \geq -p_1(\hat{c}_1)(c_1 - \hat{c}_1).$$

(3.15) follows immediately. Without loss of generality, take $c_1 > \hat{c}_1$, divide by $(c_1 - \hat{c}_1)$, and take the limit as $c_1 \rightarrow \hat{c}_1$ to obtain

$$\frac{d\pi_1(c_1|c_1)}{dc_1} = -p_1(c_1).$$

Since $\pi_1(c_1)$ is monotonic, it is Riemann integrable, thus

$$\pi_1(c_1|c_1) = \pi_1(1|1) + \int_{c_1}^1 p_1(c_1) dc_1.$$

Hence (3.13) is implied. Finally, IR clearly implies (3.17). A similar series of arguments establishes the necessity of (3.14),(3.16), and (3.18) for firm 2.

Sufficiency.

Consider firm 1. By definition of $\pi_1(\hat{c}_1|c_1)$, we have

$$\pi_1(\hat{c}_1|\hat{c}_1) = \pi_1(\hat{c}_1|c_1) - p_1(\hat{c}_1)(\hat{c}_1 - c_1).$$

Condition (3.13) implies

$$\pi_1(1|1) + \int_{\hat{c}_1}^1 p_1(s)ds = \pi_1(\hat{c}_1|c_1) - p_1(\hat{c}_1)(\hat{c}_1 - c_1),$$

or alternatively,

$$\pi_1(1|1) + \int_{c_1}^1 p_1(s)ds + \int_{\hat{c}_1}^{c_1} p_1(s)ds = \pi_1(\hat{c}_1|c_1) - p_1(\hat{c}_1)(\hat{c}_1 - c_1).$$

Simplifying, we have

$$\pi_1(c_1|c_1) = \pi_1(\hat{c}_1|c_1) - p_1(\hat{c}_1)(\hat{c}_1 - c_1) - \int_{\hat{c}_1}^{c_1} p_1(s)ds,$$

$$\pi_1(c_1|c_1) = \pi_1(\hat{c}_1|c_1) + \int_{\hat{c}_1}^{c_1} (p_1(\hat{c}_1) - p_1(s)) ds.$$

But by condition (3.15), the integral is non-negative, giving us incentive compatibility for firm 1. A similar series of arguments establishes the incentive compatibility for firm 2 using (3.14),(3.16), and (3.18).

Individual Rationality follows immediately for both firms from conditions (3.13)-(3.14) and (3.17)-(3.18). \square

Lemma 2 *The set of $\{p_i(c_1, c_2)\}_i$ which solves the modified IIR program is the same as that which solves P2 below using point-wise minimization over $\{p_i\}_i$.*

$$\min_{p_i} \left\{ p_1 \left(c_1 + \frac{F_1(c_1)}{f_1(c_1)} \right) + p_2 \left(c_2 + \frac{F_2(c_2)}{f_2(c_2)} \right) + p_0 \left(\ell(c_1, c_2) + \ell_2 \frac{F_2(c_2)}{f_2(c_2)} \right) \right\}.$$

Proof: The modified program is formally given by

$$\min_{\mathcal{M}} \int_0^1 \int_0^1 \{t_1(c_1, c_2) + t_2(c_1, c_2)\} dF_1(c_1)dF_2(c_2)$$

subject to IC and IIR.

Substituting out $t_i(c_1, c_2)$ in the objective function yields as the minimand

$$\int_0^1 \int_0^1 \{ \pi_1(c_1, c_2) + \pi_2(c_1, c_2) + p_1(c_1, c_2)c_1 \\ + p_2(c_1, c_2)c_2 + p_0(c_1, c_2)\ell(c_1, c_2) \} dF_1(c_1)dF_2(c_2).$$

We can simplify this expression by noting

$$\begin{aligned}
& \int_0^1 \int_0^1 \pi_j(c_1, c_2) dF_1(c_1) dF_2(c_2) \\
&= \int_0^1 \left\{ \pi_j(c_1, c_2) F_j(c_j) \Big|_0^1 - \int_0^1 \frac{\partial \pi_j(c_1, c_2)}{\partial c_j} dF_j(c_j) \right\} dF_{-j}(c_{-j}), \\
&= \int_0^1 \left\{ \pi_j(1, c_{-j}) - \int_0^1 \frac{\partial \pi_j(c_1, c_2)}{\partial c_j} \frac{F_j(c_j)}{f_j(c_j)} dF_j(c_j) \right\} dF_{-j}(c_{-j}), \\
&= \pi_j(1|1) - \int_0^1 \int_0^1 \left\{ \frac{\partial \pi_j(c_1, c_2)}{\partial c_j} \frac{F_j(c_j)}{f_j(c_j)} \right\} dF_1(c_1) dF_2(c_2).
\end{aligned}$$

By (3.13),(3.14),(3.17), and (3.18), it is clear that to minimize the buyer's costs, $\pi_j(1|1) = 0$, for $j = 1, 2$. Now, we substitute the above expression and use (3.13)-(3.14) to simplify to obtain the following objective function:

$$\begin{aligned}
& \int_0^1 \int_0^1 \left\{ p_1(c_1, c_2) \left(c_1 + \frac{F_1(c_1)}{f_1(c_1)} \right) + p_2(c_1, c_2) \left(c_2 + \frac{F_2(c_2)}{f_2(c_2)} \right) \right. \\
& \quad \left. + p_0(c_1, c_2) \left(\ell(c_1, c_2) + \ell_2 \frac{F_2(c_2)}{f_2(c_2)} \right) \right\} dF_1(c_1) dF_2(c_2).
\end{aligned}$$

We want to minimize this subject to conditions (3.15)-(3.16). Once done, we can substitute the resulting $p_i(c_1, c_2)$ into $\pi_j(c_1, c_2)$ and the incentive compatibility conditions to obtain the required optimal payments. Rather than minimize subject to the monotonicity constraints, we will ignore them for now, and check our solution for their satisfaction.

Choosing the optimal $p_i(c_1, c_2)$ while ignoring the monotonicity constraints for the above integrand amounts to point-wise minimization of the bracketed, expression over $\{p_i\}_i$.

To complete the Lemma, we must show the monotonicity conditions (3.15)-(3.16) hold. It is sufficient for monotonicity that $p_1(c_1, c_2)$ is non-increasing in c_1 , and that both $p_2(c_1, c_2)$ and $p_0(c_1, c_2)$ are non-increasing in c_2 . Given $\ell_1 < 1$, and given our assumptions regarding the cost distributions, this is indeed the case. \square

Finally, we show a solution to the relaxed IIR problem satisfies limited liability.

Lemma 3 *The following payments implement the optimal $\{p_i(c_1, c_2)\}_i$ for the relaxed IIR program and satisfy the limited liability constraints:*

$$\begin{aligned}
t_1(c_1, c_2) &= \int_{c_1}^1 p_1(s) ds + p_1(c_1, c_2) c_1, \\
t_2(c_1, c_2) &= \int_{c_2}^1 \{p_2(s) + p_0(s) \ell_2\} ds \\
&\quad + p_2(c_1, c_2) c_2 + p_0(c_1, c_2) \ell(c_1, c_2).
\end{aligned}$$

Proof: Substituting the above payments into (3.13) and (3.14) in the text demonstrates

the payments maintain incentive compatibility by Lemma 1. Also, the payments clearly meet the limited liability constraint, as the integrals in the above expressions are never negative for any cost realization. Finally, there do not exist any other payments with lower expected value to the buyer. This last point is evident from Lemma 2. \square

Proposition 2 in the text follows directly from Lemma 3.

Note: There are alternative payment schemes which also solve P2 but fail the limited-liability constraint. For example,

$$\begin{aligned}\tilde{t}_1(c_1, c_2) &= \int_0^1 t_1(c_1, c_2) dF_2(c_2), \\ \tilde{t}_2(c_1, c_2) &= \int_0^1 t_2(c_1, c_2) dF_1(c_1).\end{aligned}$$

These payments implement the optimal scheme, have the same incentive effect, and meet the IIR constraint with equality, but they fail the limited-liability constraints. Incentives and IIR are unchanged, because taking expectations has no effect on risk-neutral parties. However, taking expectations over the previous payments will succeed in leaving some realizations of nature with negative payoffs, violating limited liability.

Furthermore, there is another payment scheme which meets the limited liability constraints, but unlike the scheme in Lemma A, gives each firm positive payoffs in almost every state, even when the firm is not chosen to produce. Consider the following:

$$\begin{aligned}\check{t}_1(c_1, c_2) &= \int_0^1 \int_{c_1}^1 p_1(s) ds dF_2(c_2) + p_1(c_1, c_2)c_1, \\ \check{t}_2(c_1, c_2) &= \int_0^1 \int_{c_2}^1 \{p_2(s) + p_0(s)\lambda_2\} ds dF_1(c_1) \\ &\quad + p_2(c_1, c_2)c_2 + p_0(c_1, c_2)\ell(c_1, c_2).\end{aligned}$$

Proof of Proposition 4: The moral hazard problem amounts to minimizing the expected cost of the buyer's expected payments, subject to the investment constraint given in (3.10). We can now summarize the new program as P3.

$$\min_{\mathcal{M}, e} \int_0^1 \int_0^1 \left\{ \sum_{i=0}^1 p_i(c_1, c_2) \tilde{J}_i(c_1, c_2) - \psi(e) \right\} dF_1(c_1|e) dF_2(c_2), \quad (3.19)$$

subject to monotonicity in $p_1(c_1)$ and $p_2(c_2)$ and to (3.10), where the $\tilde{J}_i(c_1, c_2)$ are the virtual types for the moral hazard problem as defined in the text. As before, we ignore the monotonicity constraints and check that our solution satisfies them.

Given our assumption that $\psi(1) = \infty$, we know by (3.10) that $e < 1$. Let μ be the Lagrange multiplier associated with the constraint in (3.10) and suppose for the moment that $\mu > 0$. Minimizing the Lagrangian taking the optimal choice of $p_i(c_1, c_2)$ as given, effort is chosen such that either (3.12) is satisfied or $e = 0$. By our assumptions on F and ψ , the marginal benefit from e is positive at $e = 0$, and so we know $e \in (0, 1)$ and

(3.12) holds.

Now, given that e is optimally set at \hat{e} and given the value of $\mu > 0$, we may solve for the optimal $p_i(c_1, c_2)$. Bringing the investment constraint within the objective function yields

$$\int_0^1 \int_0^1 \left\{ \sum_{i=0}^1 p_i(c_1, c_2) \bar{J}_i(c_1, c_2) \right\} dF_1(c_1|\hat{e}) dF_2(c_2) - \mu \psi'(\hat{e}) - \psi(\hat{e}).$$

But the solution to the minimum of this expression is identical as the pointwise minimum of

$$\min_{p_i} \left\{ p_1 \left[c_1 + \frac{F_1(c_1|\hat{e})}{f_1(c_1|\hat{e})} + \mu \frac{F_{1,e}(v|\hat{e})}{f_1(c_1|\hat{e})} \right] + p_0 \left[\ell(c_1, c_2) + \ell_2 \frac{F_2(c_2)}{f_2(c_2)} \right] \right\}$$

The problem is therefore as is in the Proposition. Providing that $\mu > 0$, the virtual costs are appropriately monotone in costs so as to satisfy the additional monotonicity constraints.

Finally, we must check that $\mu > 0$. Suppose that $\mu \leq 0$ and let e^* be the choice of effort by the developer given μ . The developer will maximize its profits, not taking into account the positive externality effort has on reducing licensed costs. Note that reduced costs make $c_1 - \ell(c_1, c_2)$ decrease, and therefore produces a negative effect on developer profits. And given that the inverse hazard rate is decreasing in effort, the buyer will always prefer more effort than the developer will be willing to invest. Thus, the marginal benefit of e will be positive, and (3.12) implies that $\mu > 0$, yielding a contradiction. \square

References

- SPENCE, M., 19, "Job Market Signaling," *Quarterly Journal of Economics*, Vol. 8, pp. 355-34.
- DEMSKI, J., D. SAPPINGTON AND P. SPILLER, 1987, "Managing Supplier Switching," *Rand Journal of Economics*, pp. 77-97.
- GANSLER, JACQUES, 1989, *Affording Defense*, MIT Press.
- GROSSMAN, S. AND O. HART, 1983, "An Analysis of the Principal-Agent Problem," *Econometrica*, Vol. 51, pp. 7-45.
- HOLMSTRÖM, B., 1979, "Moral Hazard and Observability," *Bell Journal of Economics*, Vol. 10, pp 74-91.
- LAFFONT, J.-J. AND J. TIROLE, 1986, "Using Cost Observation to Regulate Firms," *Journal of Political Economy*, Vol. 94, pp. 614-641.
- LAFFONT, J.-J. AND J. TIROLE, 1988, "Repeated Auctions of Incentive Contracts, Investment and Bidding Parity with an Application to Takeovers," *Rand Journal of Economics*, Vol. 19, pp. 516-53.
- MCAFFEE, P. AND J. MCMILLAN, 1986, "Bidding for Contracts: A Principal-Agent Analysis," *Rand Journal of Economics*, Vol. 17, pp. 328-338.
- MIRRELES, J., 1971, "An Exploration in the Theory of Optimum Income Taxation," *Review of Economic Studies*, Vol. 38, pp. 15-208.
- MYERSON, R., 1981, "Optimal Auction Design," *Mathematics of Operations Research*, Vol. 6, pp. 58-73.
- NALEBUFF, B. AND J. STIGLITZ, 1983, "Information, Competition, and Markets," *American Economic Review*, Vol. 73, pp. 278-283.
- RIORDAN, M. AND D. SAPPINGTON, 1988, "Second-Sourcing," *Rand Journal of Economics*, Vol. 20, pp. 41-58.
- ROGERSON, W., 1985, "The First-Order Approach to Principal-Agent Problems," *Econometrica*, Vol. 53, pp. 1357-68.
- SAPPINGTON, D., 1983, "Limited Liability Contracts Between Principal and Agent," *Journal of Economic Theory*, Vol. 29, pp. 1-21.
- SCHARFSTEIN, D., 1988, "The Disciplinary Role of Takeovers," *Review of Economic Studies*, Vol. 55, pp. 185-199.